Winter 1989

# Isozyme and quantitative trait variation within and among natural populations of the wild soybean, Glycine soja (Sieb & Zucc)

Carol Jean Bult
*University of New Hampshire, Durham*

Order Number 9022635

# Isozyme and quantitative trait variation within and among natural populations of the wild soybean, *Glycine soja* (Sieb. & Zucc.)

Bult, Carol Jean, Ph.D.

University of New Hampshire, 1989



U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

# ISOZYME AND QUANTITATIVE TRAIT VARIATION WITHIN AND AMONG NATURAL POPULATIONS OF THE WILD SOYBEAN, GLYCINE SOJA (SIEB. & ZUCC.)

BY

CAROL JEAN BULT
B.S. George Mason University, 1984

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

in

Genetics

December, 1989

This dissertation has been examined and approved.

Dissertation director, Yun-Tzu Kiang, Professor of Plant Science and Genetics

Lincoln C. Peirce, Professor of Plant Science and Genetics

Robert T. Eckert, Associate Professor of Forest Resources and Genetics

Thomas D. Lee, Associate Professor of Botany

Anita S. Klein, Assistant Professor of Biochemistry and Genetics

Thomas M. Davis, Assistant Professor of Plant Science and Genetics

28 November 1989

Date

## DEDICATION

To my parents, Conrad and Helen Bult, who told me that I could be anything I wanted to be.

And to Christopher, for bearing with me.

## ACKNOWLEDGEMENTS

This research project was made possible possible through the intellectual and technical contributions of many people. I am grateful to them all. I, of course, bear sole responsibility for any errors or omissions in the final product.

I thank my major advisor, Dr. Yun-Tzu Kiang, for his guidance, support, generosity and encouragement throughout the course of my graduate career. Drs. Linc Peirce, Bob Eckert, Tom Lee, Anita Klein and Tom Davis served as members of my dissertation committee. I am grateful to them for their friendship and advice and for their constructive criticism of this manuscript.

On issues of statistical methodology, I consulted regularly with Owen Durgin, Will Urban, Linc Peirce and Bob Eckert. Petr Brym was invaluable in implementing many of the computer programs on the VAX that I used for data analysis. Clayton Penniman assisted in the use of the CLUSTAN computer package. Mark Hutton provided volumes of notes and examples on using SAS and SAS-GRAPH. In the lab, I benifited from the technical expertise of Yueh-Chin Chiang and Howard Doong.

Parts of this research were funded through grants from the Central University Research Fund (CURF) (1000UFRCU6D and 1000UFRCUB4) and by a UNH Dissertation Fellowship.

My graduate school tenure was enriched and enlivened by the friendship and assistance of the Plant Biology and Genetics support staff, faculty and graduate students. Thank you, all.

Karen (Raz) Rasmussen was (and is) a valued friend and colleague. I thank her especially for one particular late-night sanity break at Poco Diablo's and for all the bottles of Bully Hill wine.

Thelma Stolzenburg provided a haven from the madness with her positive attitude, typing speed, emergency monies and overflowing candy dish. Thanks for everything, Thelma.

Bill Given and Bob Parker tended the soybean jungle. Work-study students (Jim, Sue, Brenda, Linda and Katherine) helped to collect data in the greenhouse. Karen Rasmussen, David Handley, Tim Blenk, Karen Gast and Katherine Sullivan came to my rescue during the great soybean pod harvests.

I thank my parents and my sister for their love, understanding and support in this as in all things. Chris made the final months of data collection, analysis and revisions bearable and life, in general, wonderful. Clyde, the wonder cat, (1983-1989) was my faithful and excellent companion. I miss him very much.

Finally, I am grateful for the mentorship I have received over the years from Ms. Mary Bloomer and Dr. S. Laura Adamkewicz, two strong women whose enthusiasm for science and teaching inspired me to pursue a career in biology.

TABLE OF CONTENTS

APPENDICES (continued)

LIST OF TABLES

## LIST OF FIGURES

## ABSTRACT

ISOZYME AND QUANTITATIVE TRAIT VARIATION WITHIN AND AMONG NATURAL POPULATIONS OF THE WILD SOYBEAN, GLYCINE SOJA (SIEB. & ZUCC.)

by

CAROL JEAN BULT
University of New Hampshire, December, 1989

Isozymes and quantitative traits were used to explore the organization of genetic variation within and among seven natural populations of the wild soybean, Glycine soja, from Mishima, Japan. Seed enzymes were separated by horizontal starch and polyacrylamide gel electrophoresis. One-hundred and eleven individuals were scored for their genotypes at 48 gene loci resolved from 20 enzyme systems and one seed protein. Fifteen of the isozyme gene loci were polymorphic. The values for average polymorphism per locus (using the 99% criterion), average number of alleles per locus, and the proportion of heterozygotes observed were 0.140, 1.14 and 0.002, respectively. The average value for $G_{ST}$ was 0.627. The average value for Nei's genetic distance was 0.063.

Estimates of phenotypic and life-history variation in populations of wild soybean were obtained by growing seed collected in the field in two "common garden" experiments (one in 1986 and one in 1987). Of the 32 traits examined, 21 were significantly different among populations in 1986 and 18 were significantly different in 1987. In both years,

approximately 70% of the total phenotypic variation resided within populations. Canonical discriminant analysis demonstrated that the populations have diverged most significantly for those traits related to flower size, leaf shape, and yield.

The correlation between Nei's genetic distance (D) and Mahalanobis' phenotypic distance ($D^2$) was not significant in 1986 ($r_s$=0.61, p=0.14), but was significant in 1987 ($r_s$=0.85, p=0.02). This inconsistency was attributed to small sample sizes, experimental error, and phenotypic plasticity.

Variation in six reproductive yield components was examined among the seven populations of G. soja. The three early yield components (number of flower buds, flowers and immature pods per inflorescence) were significantly different among the seven populations; the three late yield components (number of mature pods and seeds per inflorescence and mature seed weight per inflorescence) were not significantly different among the populations. The rates of abortion for wild soybean ranged from 61.4% for immature pods to 3.5% for flower buds. The patterns of abortion were likely associated with resource availability.

Aspects of soybean germplasm conservation and directions for future research were discussed.

# CHAPTER I

## GENERAL BACKGROUND

### INTRODUCTION

The cultivated soybean (Glycine max (L.) Merrill) is among the world's most nutritionally concentrated and versatile foods. On average, soybean seeds contain 40% protein and 20% oil (Fehr, 1987). It has been estimated that soybeans provide one-fourth of the world's fats and oils, approximately two-thirds of the world's high protein animal feeds and three-fourths of the total world trade in protein meal (Soybean Research Advisory, 1984). Efforts to improve soybean yield, quality, yield stability, and production efficiency have been hampered by the narrow genetic base currrently used by plant breeders (Kidd, 1988; Gorman, 1983). Three major factors account for the limited nature of the soybean germplasm base. First, the majority of soybean cultivars grown in the United States, the world's largest producer and exporter of soybeans, are derived from less than 10 parental lines (Luedders, 1977; Delannay et al., 1983). Second, soybean breeders rely heavily on elite, adapted lines as sources of genetic diversity (Duvick, 1984). Third, modern pure-line breeding stocks have replaced the more genetically variable land races under cultivation in many developing countries, exacerbating the genetic erosion and vulnerability (Plucknett

et al., 1987; Duvick, 1984). Unexploited genetic diversity exists in the current G. max germplasm (Gorman, 1983); however, additional sources of gene diversity are urgently needed to meet short term crises and long term breeding goals (Duvick, 1984; Kidd, 1988).

Natural populations of a crop species are particularly important for germplasm conservation programs. These populations are genetically adapted to heterogeneous environments and often have served as excellent sources of novel genes and gene combinations for disease resistance, physiological adaptations for drought, heat and salt tolerance, and quantitative traits of economic importance (Harlan, 1976; Frankel and Soule, 1981; Williams, 1982; Plucknett et al., 1987). Natural populations of the wild soybean (Glycine soja Sieb. & Zucc.) can serve as essential reservoirs of genetic diversity for soybean breeding programs (Kiang et al., 1987). G. soja generally is regarded as the wild progenitor to G. max, and together the two species form the soybean gene pool (Kiang et al., 1987). A meager 30% of the estimated available genetic diversity in wild soybean populations worldwide has been collected and characterized for germplasm conservation (Plucknett et al., 1987). The need for further evaluation is critical for wild soybean as many of its natural habitats are being destroyed due to expansion of urban areas and modern agricultural practices (Plucknett et al., 1987; Nevo, 1987).

## Use of wild soybean germplasm in breeding programs

The potential value of *G. soja* in soybean breeding programs has been demonstrated. Compared to the cultigen, *G. soja* has a higher DNA content (in picograms) per cell and per chromosome (Yamamoto and Nagato, 1984), greater percentage of seed protein (Kaizuma and Fukui, 1974) and greater total phosphorous, zinc and calcium content (Raboy et al., 1984). Wild soybean germplasm has been used successfully as a source of genes for disease resistance (Ram et al., 1984), small seed size (Fehr, 1987) and increased seed protein (Kaizuma et al., 1980; Erickson et al., 1981).

*G. soja* has been used only on a limited basis in improvement programs thus far because the species has a number of agronomically undesirable traits, including hardseededness, shattering of pods, retention of leaves and petioles at maturity and asynchronous flowering and seed maturity which often are expressed in interspecific hybrid progeny (Ertl and Fehr, 1985; Carpenter and Fehr, 1986). *G. soja* germplasm has not shown the same value in yield improvement as has been acheived with wild progenitors of other crop species (Lawrence and Frey, 1975; Harlan, 1976). Ertl and Fehr (1985) failed to recover transgressive segregants for yield in six backcross generations between two *G. max* cultivars and two *G. soja* plant introductions (PIs). However, the amount of variation for yield, height and

lodging resistance was markedly increased, suggesting that exotic germplasm may have some long term benefit for improvement of quantitative traits in soybean.

Genetic variation in G. max and G. soja

Compared to other major crops, such as wheat, barley and maize, little is known of the organization of genetic variation in natural populations of the wild soybean or of the evolutionary relationship between the two species (Kiang and Gorman, 1983; Kiang and Chiang, 1989). Kiang et al. (1987) characterized the genetic variation via electrophoresis in G. max and G. soja accessions from China, Japan, Korea, Taiwan, USA and USSR. These accessions represent collections of a few plants from diverse geographical areas and are not necessarily representative of a natural population of soybeans. Kiang and coworkers found genetic variation to be fairly evenly distributed geographically for G. max accessions, presumably due to the extensive human influence on the distribution, evolution and cultivation of this species. The levels of genetic variation among the accessions were comparable to those of other cultivated, self-pollinating species (Kiang et al., 1987). In contrast, genetic variation in G. soja was higher than expected for selfing annuals, was heterogeneous in its distribution and contained alleles not found in the cultivated soybean germplams (Kiang et al., 1987). Clearly,

wild soybeans are a potentially valuable source of genetic variation to broaden the soybean germplasm base.

Studies of genetic variation in natural populations of wild soybeans to date have focused on soybean populations collected over considerable geographic distances (Zhi-ang and Hong-xin, 1985; Chiang, 1985). As pointed out by Loveless and Hamrick (1984), sampling over "different geographical scales" is needed to enhance understanding of genetic structure in plant populations; thus, the present study will focus on genetic variation within and among local natural populations of wild soybean.

## The organism

Soybeans are members of the Leguminosae, subfamily Papilionoideae, tribe Phaseolae. The genus, Glycine Willd. currently is divided into two subgenera, Glycine and Soja (Hadley and Hymowitz, 1987). Subgenus Glycine contains the seven perennial soybean species; subgenus Soja contains the cultigen, [G. max (L.) Merr.] and the wild soybean [G. soja (Sieb. & Zucc.)]. Both species are predominantly self-fertilizing annuals with chromosome number 2N=40 (Hymowitz, 1970).

G. soja is the presumed wild progenitor to G. max. It is thought to have originated in northeast China and currently is distributed throughout China, Japan (except Ryukyu Islands), the Korean peninsula, the eastern Siberian regions

of the USSR and Taiwan (Hymowitz, 1970; Kiang et al., 1987). Although occasionally harvested for forage, G. soja is not grown as a cultivated crop (Hymowitz, 1970). The wild soybean plants grow in fields, under hedgerows and along roadsides and riverbanks. They have vine-like stems and climb on and among neighboring plants. Their leaves are pinnately trifoliate, and the leaflets may be narrowly lanceolate, ovate or oblong-elliptic. Pubescence on pods, stems and leaves generally is tawny and is strigose to hirsute. Pods are short (1 to 4 cms), black, and contain small dark brown to black oval seeds. The flowers are "typically papilionaceous with a tubular calyx of five unequal sepal lobes, and a five-parted corolla consisting of a posterior banner petal two lateral wing petals, and two anterior keel petals in contact with each other but not fused" (Carlson and Lersten, 1987). The flowers are purple or, rarely, white. The flower morphology is identical to that of G. max although the flowers of G. soja generally are smaller in size (personal observation).

G. max and G. soja are interfertile, and zones of natural hybrids between the two have been found in nature, although no hybrid swarms have been observed (Kiang and Chiang, 1989). Although G.max and G. soja are quite distinct morphologically, evidence from chromosomes (Hadley and Hymowitz, 1987; Palmer et al., 1987), isozyme analyses (Kiang and Gorman, 1983) and restriction endonuclease analyses of

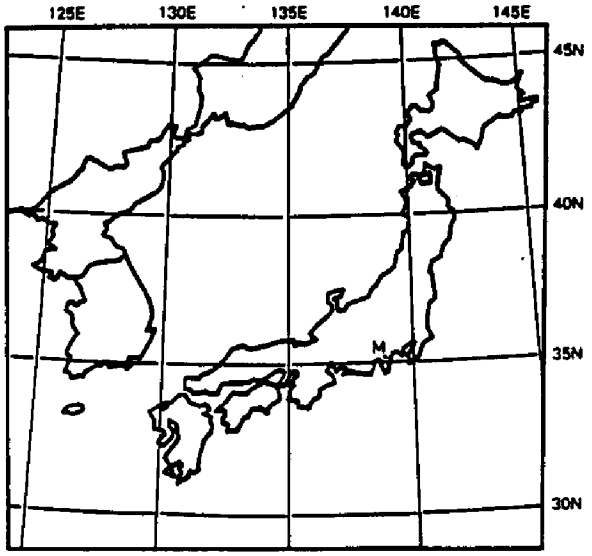
ribosomal RNA (Doyle and Beachy, 1985; Doyle, 1988), mitochondrial DNA (Sisson et al., 1978) and chloroplast DNA (Shoemaker, et al., 1986) indicates that the two are as similar as plants within the same species (Kiang et al., 1987). Shoemaker et al. (1986) have suggested that they be considered subspecies.

Soybeans may be ancient tetraploids with diploidized genomes (Hadley and Hymowitz, 1987). Based on cytological evidence within the tribe, Hadley and Hymowitz (1987) have hypothesized that the chromosome number of Glycine genus is derived from diploid ancestors with a base chromosome number of 11, which, through aneuploid reduction, became 10. Subsequent polyploidization produced somatic chromosome numbers of 40 or sometimes 80.

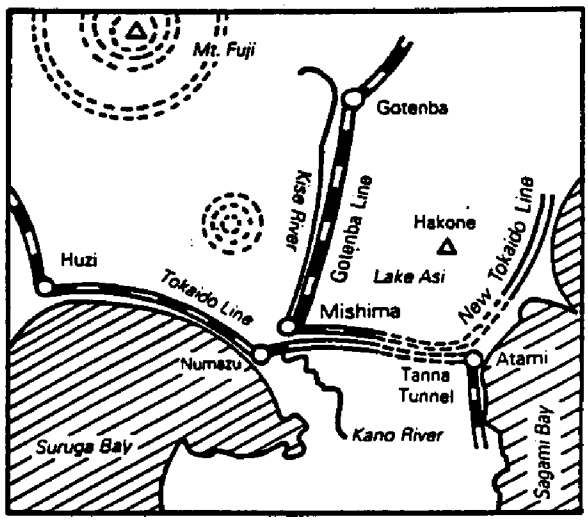
Seed source

The G. soja seed material used in this research was a generous gift of Drs. H.I. Oka and H. Morishima (of the National Institute of Genetics, Mishima, Japan).

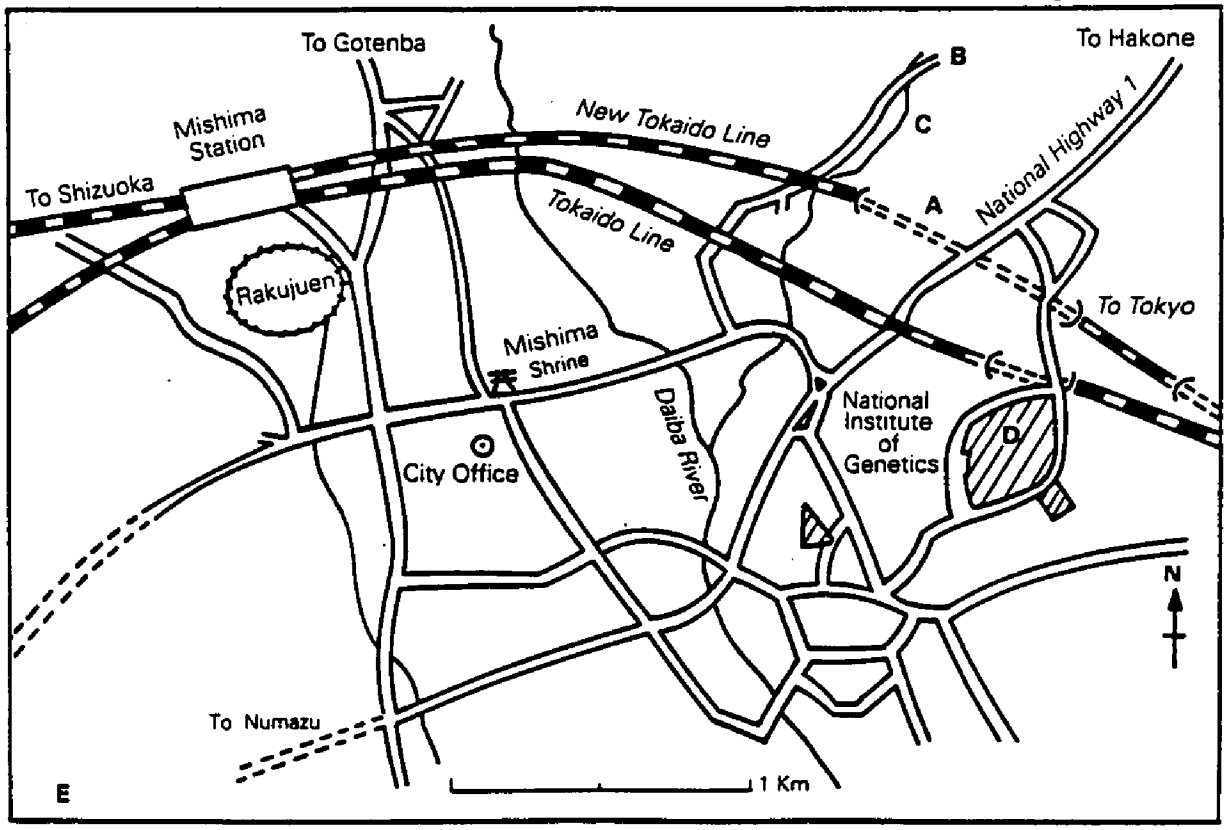
The seeds were hand-collected from 111 individual plants over seven sites in and around the city of Mishima, Japan in 1982 and 1983 (Figure 1; Table 1). The collection sites varied with respect to kinds and percentages of co-habitating plants, degree of site disturbance, soil type, soil moisture, shading and slope. Five of the populations (1, 2, 3, 5 and 6) were collected from the western foot of the




1a. Location of Mishima, Japan (M).




1b. Area map showing the relation of Mishima to surrounding cities.




1c. Detail map of Mishima showing regions of soybean seed collections (A-E).

Figure 1a-c. Maps showing the location of Mishima, Japan and seed collection sites for the Glycine soja populations used in this study. Populations 1 and 5 were from region A; 2 and 6 were from region B; 3 was from region C; 4 was from region D; and 7 was from region E (see Table 1 for further details).

Table 1. Collection site descriptions for seven natural populations of Glycine soja from Mishima, Japan. N is the number of plants from which seeds were collected at each site.

| Population | Year collected | N | Site description |
|---|---|---|---|
| 1. Hatsunedai | 1982 | 13 | A roadside strip along a rice field. |
| 2. Kanodanchi | 1982 | 36 | A roadside strip used for ornamental plants. |
| 3. Nishi-asahigaoka | 1982 | 10 | A strip of land between the road and shrubs. |
| 4. Yata | 1982 | 30 | An unused field on the campus of the National Institute of Genetics. |
| 5. Asahigaoka-1 | 1983 | 10 | A ruderal field left for housing. |
| 6. Asahigaoka-2 | 1983 | 9 | A roadside slope. |
| 7. Kakitagawa | 1983 | 3 | A moist site on a river terrace. |

now inactive Hakone volcano. One population (4) was collected from an unused field on the campus of the National Institute of Genetics and one (7) was collected from a river terrace just outside Mishima's city limits. The geographic distances between collection sites ranged from less than 1 km to over 3 km. The dimensions of the sampling areas ranged from 5 X 1 meters to 100 X 2.5 meters. Descriptions of the sites and collectors' comments about the collections are given in Appendix I.

## Objectives of this research

The major goal of this research was to characterize isozyme and quantitative trait variation within and among local, natural populations of the wild soybean. Specific objectives included the following:

1) To investigate the organization of isozyme variation within and among local, natural populations of G. soja from central Japan (Mishima);

2) To quantify the variation in a number of morphological, agronomic and phenological traits within and among the soybean populations;

3) To examine the degree of congruence between isozymes and quantitative traits in describing population differentiation;

4) To explore plasticity in seed yield components within and among G. soja populations; and,

5) To examine levels and patterns of abortion of flower buds, flowers, pods and seeds in the wild soybean.

# CHAPTER II

## ISOZYME VARIATION WITHIN AND AMONG NATURAL POPULATIONS OF *GLYCINE SOJA*

## INTRODUCTION

A population is a dynamic unit which must be genetically flexible in the face of environmental heterogeneity if it is to survive. Levels and patterns of genetic variation are essential elements conditioning the evolutionary potential of a population (Wright, 1951; Dobzhansky, 1970). The amount of genetic variation determines the capacity of a population to adapt to its local environment, while the distribution of genetic variation conditions the "modes and tempos of evolutionary change" (Soltis and Soltis, 1987). Population structure and differentiation are the result of the joint action of mutation, migration, selection, drift and gene flow operating in the historical and biological context of a particular species or population (Ayala, 1982a; Soltis and Soltis, 1987). In general gene flow homogenizes populations, whereas mutation and selection promote differentiation.

### Estimating genetic variation via protein polymorphisms

In the 1950's and 60's, electrophoretic detection of isozyme variants became the standard molecular approach in experimental population genetics (Lewontin and Hubby, 1966;

Hubby and Lewontin, 1966). Since that time, gel electrophoresis has been used extensively to explore genetic variation in literally thousands of species of organisms (Gottlieb, 1981). The theory and practice of electrophoresis are straightforward. Proteins migrate in an electric field through an inert gel matrix according to their overall charge, molecular weight and, to some degree, conformation. These factors are conditioned by the amino acid sequence of a protein. If the amino acid sequence changes due to a mutation, the relative electrophoretic mobility of the protein also may change. Electrophoretic mobility variants, therefore, indirectly reflect genic diversity. To visualize areas of enzyme-activity on a gel, a gel slice is placed in a buffer solution containing specific substrates, cofactors, coenzymes and a dye which is bound, or colorized, where the substrate is acted upon (Lewontin, 1974).

The advantages and pitfalls of gel electrophoresis have been discussed in great detail (Lewontin, 1974; Gottlieb, 1971, 1981; Brown and Weir, 1983). The advantages include the following: 1) expression of alleles usually is codominant and generally is not affected by environment; 2) a sample of loci can be chosen based on availability of a reliable enzyme-activity assay, irrespective of variability at the locus in question; 3) substitution of an allele at one locus is distinguishable from that at another; and 4) allelic differences are detected as mobility differences and

are not dependent on the functional role of the enzyme or its overall level of variation (Brown and Weir, 1983). In addition, electrophoretic results are highly repeatable, and a large number of samples can be assayed relatively quickly and inexpensively.

The major disadvantages of electrophoresis are 1) not every amino acid substitution results in an electrophoretically distinguishable mobility variant, 2) not all enzymes can be detected with colorized products and 3) only 10% of the eukaryotic genome codes for soluble protein products (Gottlieb, 1981). Thus, non-translatable regions of DNA, membrane bound enzymes and many regulatory genes cannot be resolved and evaluated using this method.

## Genetic diversity estimates

The estimates of genetic diversity based on enzyme variation are typically expressed in terms of the average number of alleles per locus, polymorphism and expected heterozygosity (Brown and Weir, 1983). The range and magnitude of these estimates are dependent on the number and kinds of enzyme loci and number of individuals examined (Lewontin, 1974). Comparisons of these statistics among different studies and species must, therefore, be made with caution.

The number of alleles per locus (A) estimates allelic richness, one component of genetic diversity (Brown and Weir, 1983). It suffers somewhat as a measure of diversity

because it gives equal weight to all alleles, without regard to physiological or biological significance of a particular allelic variant (Brown and Weir, 1983). It also is inflated when rare alleles are included in an analysis (Nei, 1987).

A polymorphic locus is one at which the most common allele has a frequency of not greater than 0.95 or 0.99. These frequencies are chosen arbitrarily and are intended to exclude the most rare alleles from the statistic. Polymorphism (P) is, at best, a rough estimation of genetic variation, because a slightly polymorphic locus (i.e. one with few alleles) is weighted as heavily as a very polymorphic one (i.e. one with many alleles) (Ayala, 1982a).

Gene diversity (also called average or expected heterozygosity) is a less arbitrary and more accurate measure of genetic variation, because it is based on gene frequencies (Nei, 1973). Its calculation is discussed in Materials and Methods (page 26).

## Genetic variation within and among populations

Numerous surveys of biochemical and morphological variation in plant populations have demonstrated that individual genotypes are not randomly distributed, but are spatially and temporally structured (Bradshaw, 1972; Nevo, 1978; Kahler et al., 1980; Gottlieb, 1981). Empirical evidence from these analyses has resulted in specific expectations and generalization for the genetic structure in

natural populations of plants (Nevo,1978; Brown, 1979; Hamrick et al., 1979; Loveless and Hamrick, 1984).

Breeding system is a major determinant of genetic structure in plant populations (Brown, 1979; Hamrick, 1982). Compared to populations of outcrossers, predominantly inbreeding populations typically are less heterozygous, less diverse genetically, highly differentiated, and are characterized by multilocus associations (Hamrick et al., 1979; Brown, 1979; Loveless and Hamrick, 1984). Inbreeding per se is not expected to affect such parameters as the number of polymorphic loci or the number of alleles per locus (Schoen, 1982). Rather, inbreeding increases the probability that individuals in a population possess alleles that are identical by descent, reduces recombination and maintains gametic phase disequilibrium (Allard et al., 1968; Loveless and Hamrick, 1984; Parker, 1988). Therefore, while populations of inbreeders tend to have less within-population genetic diversity, they often maintain large stores of genetic variation among populations (Brown, 1979; Allard et al., 1968). The extent of genetic differentiation will depend on the interactions of such factors as effective population size (Slatkin, 1987), the patterns of spatial and temporal environmental variation (Hedrick, 1986), and mechanisms of pollen and seed dispersal (Allard et al., 1968).

In this chapter, I will explore the amounts and

distributions of isozyme variation within and between seven local, natural populations of the wild soybean, G. soja, from Mishima, Japan.

## MATERIALS AND METHODS

### A. Seed material

Ten to thirty seeds were hand-collected from 111 wild soybean plants on seven sites in Mishima, Japan during 1982 and 1983 (Figure 1). Descriptions of the collection sites are given in Appendix I. Seeds were stored at UNH at 0° C for two years prior to this study.

### B. Gel electrophoresis

The electrophoretic techniques used were modeled after those of Gorman (1983), Gorman and Kiang (1977), Kiang and Gorman (1983), Chiang (1985), and Doong (1986) with some modifications (see Bult et al.,.1989). Twenty enzymes and one soybean seed protein were examined in this study (Table 2). Initially, seven seeds were screened for each plant to determine its genotype. Because the number of seeds collected from the field was small, some of the seeds used in the initial screening were progeny of plants grown from field-collected seed. At least five of the seeds used for electrophoresis were from the original parent. If the bands were not resolved sufficiently or a heterozygote was detected, additional seeds (progenies of field-collected seed) were examined to confirm the genotype.

Table 2. Isozymes examined for G. soja seed.

| | Enzyme/Protein | EC No. | Type |
|---|---|---|---|
| 1. | Acid phosphatase | 3.1.3.2 | Hydrolase |
| 2. | Aconitase | 4.2.1.3 | Lyase |
| 3. | Alcohol dehydrogenase | 1.1.1.1 | Oxioreductase |
| 4. | Beta-amylase | 3.2.1.2 | Hydrolase |
| 5. | Diaphorase | 1.6.2.2 | Oxioreductase |
| 6. | Endopeptidase | 3.4.?.? | Hydrolase |
| 7. | Esterase | 3.1.1.1 | Hydrolase |
| 8. | Fluorescent esterase | 3.1.1.2 | Hydrolase |
| 9. | Glucose 6-phosphate dehydrogenase | 1.1.1.49 | Oxioreductase |
| 10. | Glutamate oxaloacetic transaminase | 2.6.1.1 | Transferase |
| 11. | Isocitrate dehydrogenase (NADP-active) | 1.1.1.42 | Oxioreductase |
| 12. | Kunitz trypsin inhibitor | ******* | Seed protein |
| 13. | Leucine aminopeptidase | 3.4.11.1 | Hydrolase |
| 14. | Malate dehydrogenase | 1.1.1.37 | Oxioreductase |
| 15. | Mannose 6-phosphate isomerase | 5.3.1.8 | Isomerase |
| 16. | Peroxidase | 1.11.1.7 | Oxioreductase |
| 17. | 6-Phosphogluconate dehydrogenase | 1.1.1.43 | Oxioreductase |
| 18. | Phosphoglucose isomerase | 5.3.1.9 | Isomerase |
| 19. | Phosphoglucomutase | 2.7.5.3 | Transferase |
| 20. | Shikimate dehydrogenase | 1.1.1.25 | Oxioreductase |
| 21. | Urease | 3.5.1.5 | Hydrolase |

1. Gel preparation. Three types of electrophoretic gels were used: acrylamide, starch and a mixture of the two. Details of gel preparation and buffer formulations are given in Appendix III. By using different kinds, combinations and concentrations of materials, gel porosity could be changed to enhance the resolution of enzyme-activity banding patterns. The cofactors, NADP and NAD, were added to some of the gels to enhance staining intensity of the bands (See Appendix III). The acrylamide/starch concentrations used in this study were 1) 12.5% starch, 2) 7% acrylamide, 3) 9% acrylamide, 4) 7% acrylamide + 2% starch, and 5) 6% acrylamide + 4% starch. For gels containing acrylamide, the total amount of gelling agent was prepared as 95% acrylamide and 5% N,N'-methylene-bis-acrylamide. Gel polymerization catalysts, ammonium persulfate (APS) and N,N,N',N'-tetramethylethylene diamine (TEMED), were 0.1% (w/v) and 0.2% (v/v) of the total volume of the gel buffer, respectively.

Gel molds were made of PVC (polyvinylchlorine) board and were of three dimensions (length X width X depth): 20.5 X 17.8 X 0.3 cm (one layer), 20.5 X 17.8 X 0.6 cm (two layer) and 20.5 X 17.8 X 0.9 cm (three layer). By using the deeper molds, multiple gel slices could be removed and stained for different enzymes having similar electrophoretic mobilities, thereby conserving materials, time and space.

2. Sample preparation. Whole dry cotyledons were used for electrophoresis. Once the seed coats were removed, the seeds were placed in individual wells on a grinding tray (several rows of shallow wells drilled into PVC board) and soaked in 2-4 drops (depending on the amount of tissue) of 0.005 M L-histidine (HCl) (pH 7.0) buffer for at least 8 hours. To screen for urease isozymes, a small portion of the cotyledon was instead soaked in a extraction buffer consisting of 60 mM Tris-HCl (pH 8.2) with 15 mM $CaCl_2$, 390 mM sucrose and 10 mM dithiothreitol (DTT) (Kloth et al., 1987). The DTT was added to the extraction buffer immediately before use. This extraction buffer was used to prevent non-genetically controlled aggregation of the urease polypeptide subunit (Polacco and Havir, 1979). All seeds were homogenized by hand using glass rods with rounded, smooth ends. A 1 X 1 cm square of lens paper then was placed on top of each sample well. An appropriate size wick cut from bibulous paper was placed on the filter paper to absorb the sample slurry. The size of the wicks used in each gel depended on gel thickness and sample number per gel.

3. Loading the gel. A scalpel, held vertically against a straightedge, was inserted into the gel approximately 3 cm from the edge of the gel mold and drawn across and through the gel to form a sample well. Alternatively, individual wells were formed by inserting into the gel a ruler to which

26 razor blade fragments (1 X 0.5 cm) were attached and arranged evenly, like teeth on a comb. Using forceps, wicks were removed from the wells of the grinding tray one at a time, blotted lightly on an absorbent tissue and inserted perpendicularly into the sample well in the gel. The wicks were placed evenly along the sample well using the markings on a straightedge as guides. At least 5 mm of space was left on either end of the line of wicks to eliminate edge effects. When all the wicks had been inserted, the two sections of the gel were pressed lightly together along the suture line to remove air gaps and to ensure good contact between the wicks and the gel.

4. Electrophoresis. Rubbermaid® plastic trays (22.5 X 8 X 5 cm) served as electrode buffer reservoirs. The trays were modified by drilling a small hole into one end of the tray, approximately 1 cm from the top. A female portion of a banana plug was fitted into this hole. One end of a 22 cm length of platinum wire was wound once or twice around the metal portion of the banana plug on the inside of the tray. Several dollops of silicon gel were used to secure the wire to the bottom of the tray and to seal any small gaps around the female banana plug fitting. The trays were connected to an ISCO Model 493 Electrophoresis Power Supply unit by electrical leads fitted with male banana plugs.

A gel, loaded with samples, was placed on a block of ice

between two electrode buffer reservoirs, each containing 100 ml of a 0.13 M Tris-citrate buffer (pH 7.0). One end of a cellulose sponge (17.5 X 19 X 0.002 cm), moistened with electrode buffer, was allowed to overlap the gel 0.5 cm from the line of wicks. The other end of the sponge was placed in the electrode buffer reservoir. On the opposite end of the gel, the sponge was placed so that it overlapped the gel by 3 cm. A schematic of this type of electrophoresis apparatus is shown below (Figure 2).

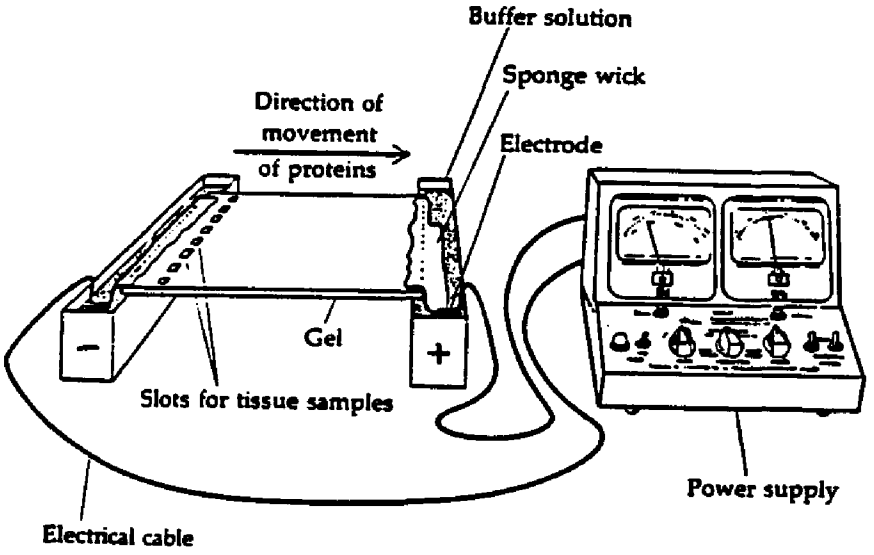



Figure 2. An example of a set up for slab gel electrophoresis. From: Hartl, D.L. 1980. Principles of Population Genetics. Sinauer Assoc., Inc. p. 73.

The tip of a pair of forceps was dipped into a 1% methylene blue solution and then inserted into the gel 1 cm from the edge of the gel mold, just below the cellulose sponge opposite the line of samples. The mobility of the dye relative to the mobility of the enzyme-activity bands was used to calculate $R_f$ values.

A piece of plastic wrap was placed over the gel and sponges. A piece of PVC board then was placed on top of the plastic wrap to ensure even contact between the sponges and the gel surface. A second, smaller block of reusable "blue ice" was placed on top of the PVC board to maintain a constant temperature of 4° C throughout electrophoresis. Gels were electrophoresed at constant voltage (See Appendix III for voltage settings).

5. Enzyme-activity staining. Following electrophoresis, the gels were sliced horizontally into several slices depending on the thickness of the gel. The slicing surface was a piece of PVC board with two raised edges (3 mm). The gel slicing tool was a modified coping saw with a 0.009 mm guitar string stretched tightly across one end. For staining ACO, LAP, MPI, and IDH, both the surface and the bottom layers were discarded. The remaining center slice, which was about 3 mm thick, was stained for the enzyme. For other enzymes, only the surface layer was discarded. Each gel slice was placed in a Pyrex baking dish (or similar container) containing a stain solution for a specific enzyme or protein. The staining recipes were adapted from a number of published sources including, Brewbaker et al. (1968), Shaw and Prasad (1969), O'Malley et al. (1980), Cardy et al. (1980), Conkle et al. (1982), Vallejos (1983) and others.

The specific protein and enzyme staining protocols and stain buffer formulations used in this study are given in Appendix III.

C. Genetic interpretation of zymogram patterns

Not every band on a gel is the direct product of a gene or genes. Bands may result from interaction of proteins with buffers, molecular instability, proteins binding with charged factors, post-translational modification, contamination, and formation of a polymer series (Gottleib, 1981; Moss, 1982). True isozymes are variant forms of enzymes whose synthesis is conditioned by distinct sequences of DNA including, protein variants produced by independent gene loci (isozymes), allelic protein variants (allozymes) and non-convalently bound polypeptide chains (heteropolymers). Transient variant forms of enzymes which originate by post-genetic modification of a single polypeptide chain, and covalently modified or conformationally different forms are not true isozymes (Dixon and Webb, 1979). Varying electrophoretic conditions can help in distinguishing between artifacts and true enzymatic activity (Ramshaw et al., 1979; Ayala, 1982b).

The genetic basis of zymogram patterns must be understood before they can be used reliably to estimate parameters of genetic diversity. Genetic models are generated by controlled, hand-pollinated crosses between plants showing

recognized electrophoretic mobility variants. The $F_1$, $F_2$ and $F_3$ progenies of these crosses then are examined for the segregation of the electrophoretic variants and tested for conformation to expected genetic ratios.

D. Quantification of genetic variation

BIOSYS-1 (Swofford and Selander, 1981), a FORTRAN computer program mounted on the UNH VAX computer was used to calculate estimates of genetic diversity and population differentiation based on the isozyme data.

1. Estimates of genetic variation within populations. The following genetic diversity parameters estimated within-population variation:

i) Allele frequencies. Allele frequencies were calculated based on the genotypes determined via electrophoresis.

ii) Polymorphism. Polymorphism (P) was reported at both the 95 and 99% levels. The percent polymorphism was calculated as,

P = number of polymorphic loci/total number of loci examined.

iii) Allelic richness. The average number of alleles per locus (A) was calculated as,

A = Total number of alleles/ Total number of loci examined.

The total number of loci was an estimate, since in all instances the exact number of genes coding for the isozymes was not known (e.g. no electrophoretic variants had been detected or no formal genetic analysis had been performed).

iv) Heterozygosity. The average number of heterozygous individuals per locus was calculated using the algorithm,

$$H_{EXP} = 1 - \sum_{i=1}^{n} x_i^2,$$

where $H_{EXP}$ is the expected heterozygosity at a locus, n is the number of alleles, and $x_i^2$ is the frequency of homozygotes for the ith allele (Nei, 1978).

Mean expected heterozygosity over all loci was calculated as,

$$\bar{H}_{EXP} = \sum_{i=1}^{n} H_i / n$$

where $H_i$ is the heterozygosity at the ith locus and n is the number of loci examined.

2. Structuring of genetic diversity among populations. Four methods were employed to analyze the distributions of genetic variation among populations.

i) Contingency Chi-square of allele frequencies. A heterogeneity contingency Chi-square analysis was used to

test the null hypothesis that allele frequencies were homogeneous across all 7 populations (Workman and Niswander, 1970).

ii) **Partitioning genic diversity.** Nei (1973) developed statistical methods for examining how genetic variation is organized within and between populations. Nei's methods are similar to those developed by Wright (1951, 1964, 1978). Wright's F-statistics emphasize the relative reduction in heterozygosity associated with various levels of organization, whereas the genetic diversity statistics of Nei emphasize the way total genic diversity is partitioned (Greenlee and Rai, 1986). According to Nei (1973), genetic diversity in a population or group can be partitioned into within- and among-population components using the formula

$$H_T = H_S + D_{ST},$$

where $H_T$ is the total gene diversity of all populations, $H_S$ is the average gene diversity within populations and $D_{ST}$ is the average gene diversity among populations. $G_{ST}$ is the coefficient of differentiation among populations and is calculated as $D_{ST}/H_T$.

iii) **Genetic distance among populations.** Genetic distance (D) estimates the number of allelic substitutions per locus that have occurred in the separate evolution of two populations (Nei, 1972). Genetic distance may range from

zero to infinity. Genetic identity (I) estimates the proportion of alleles identical in two populations. The values for I may range from zero to one. Genetic distance and identity are related by the expression

$$D = -\ln(I).$$

Nei's genetic identity (I) between two populations, x and y is calculated as

$$j_x = \sum_{i=1}^{n} x_i^2,$$

$$j_y = \sum_{i=1}^{n} y_i^2,$$

$$j_{xy} = \sum_{i=1}^{n} x_i y_i,$$

$$I = J_{xy} / (J_x J_y)^{1/2};$$

where n is the number of alleles; $x_i$ and $y_i$ are the frequencies of the ith alleles in populations x and y, respectively; $j_x$ and $j_y$ are the probability of two alleles being identical by descent in populations x and y, respectively; $j_{xy}$ is the probability of an allele from population x and y being identical by descent; and $J_{xy}$, $J_x$, $J_y$ are the arithmetic means, over all loci, of $j_{xy}$, $j_x$, and $j_y$, respectively.

**iv) Canonical discriminant analysis of allele frequencies.**

Relationships among the seven populations of wild soybean were explored further using canonical discriminant analysis (CDA). CDA is a multivariate data reduction method. "The discriminant model is set up so that the first discriminant function maximally separates groups. Then, a second dimension, uncorrelated with the first, is found that best separates groups on the basis of information not accounted for by the first discriminant function" (Norusis, 1985). The analysis was performed using the PROC CANDISC subroutine in SAS (SAS Insititute, 1985b). Frequencies for the 31 alleles from 15 variable isozyme loci were used as the data matrix. From this matrix, a variance-covariance matrix was derived from which canonical variates were extracted.

The assumptions for a discriminant analysis are that the data follow a multivariate normal distribution and that the variance-covariance matrices are homogeneous. While it is unlikely that the allele frequencies met these assumptions, the analysis was used as an exploratory tool to detect patterns of population differentiation, not to determine if the differences among the populations were statistically significant.

The significance of the canonical variates was tested using Wilk's lambda, the default test criterion in SAS. The null hypothesis was that the current canonical correlation and all correlations which followed were equal to zero (SAS

Institute, 1985b). The distribution of the test criterion was transformed into a statistic with an approximate F-distribution (SAS Institute, 1985b).

## RESULTS

### A. Genetic models for isozyme variants

The following section describes the current genetic control models for the isozymes used in this study, their hypothesized (or known) biological significance, subunit structure and the results for the populations examined in the present study. All bands are anodal unless otherwise noted. Alleles are alphabetized consecutively from the anode. Genetic symbols for loci and alleles were assigned in accordance with the "Rules for Genetic Symbols" approved by the Soybean Genetics Committee and published in the Soybean Genetics Newsletter (Vol. 14:8-10, 1987).

1. Acid phosphatase (AP). Acid phosphatases catalyze the hydrolysis of monoesters of phosphoric acid (Dixon and Webb, 1979). The in vivo significance of these enzymes in plants is unknown. They are visualized in vitro using the synthetic substrate, alpha-naphthyl acid phosphate. To date, all acid phosphatases examined for plants have been monomeric in structure (Manchenko, 1988).

Three anodal electrophoretic variants were observed in G. max (Gorman and Kiang, 1977; Kiang, 1987) and were determined to be inherited as codominant alleles at a single nuclear locus (Ap-a, $R_F$=0.45; Ap-b, $R_F$=0.48; and Ap-c,

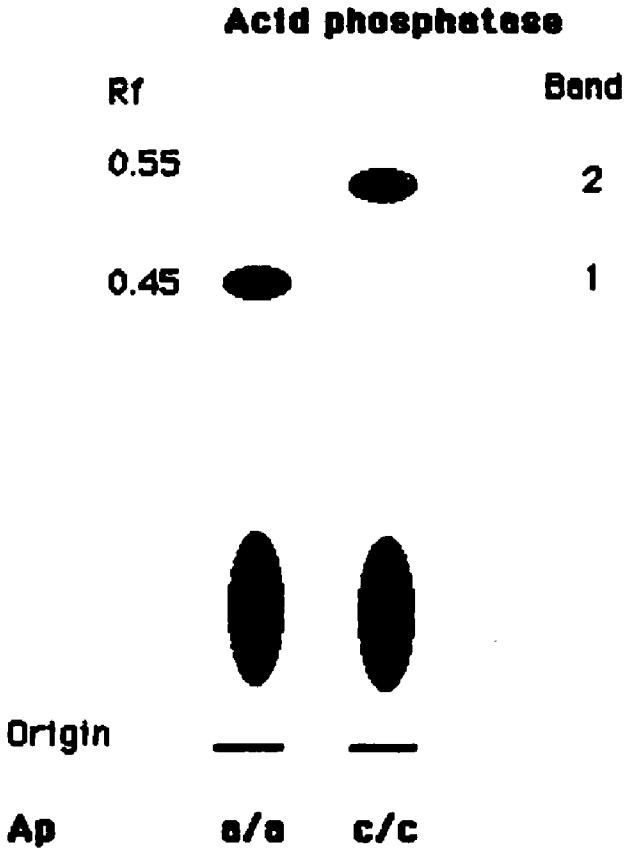



Figure 3. AP banding patterns for G. soja seed tissue.

$R_F$=0.53) (Hildebrand et al., 1980). A fourth variant was detected only in G. soja accessions and was determined to be another codominant allele (Ap-d, $R_F$=0.57) at the Ap locus (Chiang, 1985).

The Ap-a and Ap-c alleles were observed in the present study (Figure 3). A smear, consisting of an unknown number of bands of enzyme activity, was consistently observed near the origin (Figure 3). The genetic basis of these bands is not yet understood.

2. Aconitase (ACO). Aconitase isozymes mediate the interconversion of three tricarboxylic acids (citrate, cis-aconitate and isocitrate) in the Kreb's cycle (Lehninger, 1982). All ACO isozymes detected in soybeans to date have been monomeric in structure (Chiang, 1985).

Homozygotes display up to six bands for dry cotyledon tissue from G. max and G. soja and at least five loci are hypothesized to account for the banding patterns (Chiang, 1985; Doong and Kiang, 1987a; Griffin and Palmer, 1987). Two mobility variants were observed for Aco1: Aco1-a ($R_F$=0.21) and Aco1-b ($R_F$= 0.26). Four codominant alleles were observed in soybean seed for Aco2: Aco2-a ($R_F$=0.32), -b ($R_F$=0.33), -c ($R_F$=0.37) and -d ($R_F$=0.43). Two 2-band zymograms were observed for the products of Aco3. Aco3-a produces bands at $R_F$= 0.45 and 0.50, while Aco3-b produces bands at $R_F$= 0.50 and 0.56. Chiang (1985) speculated that the $R_F$=0.50 band

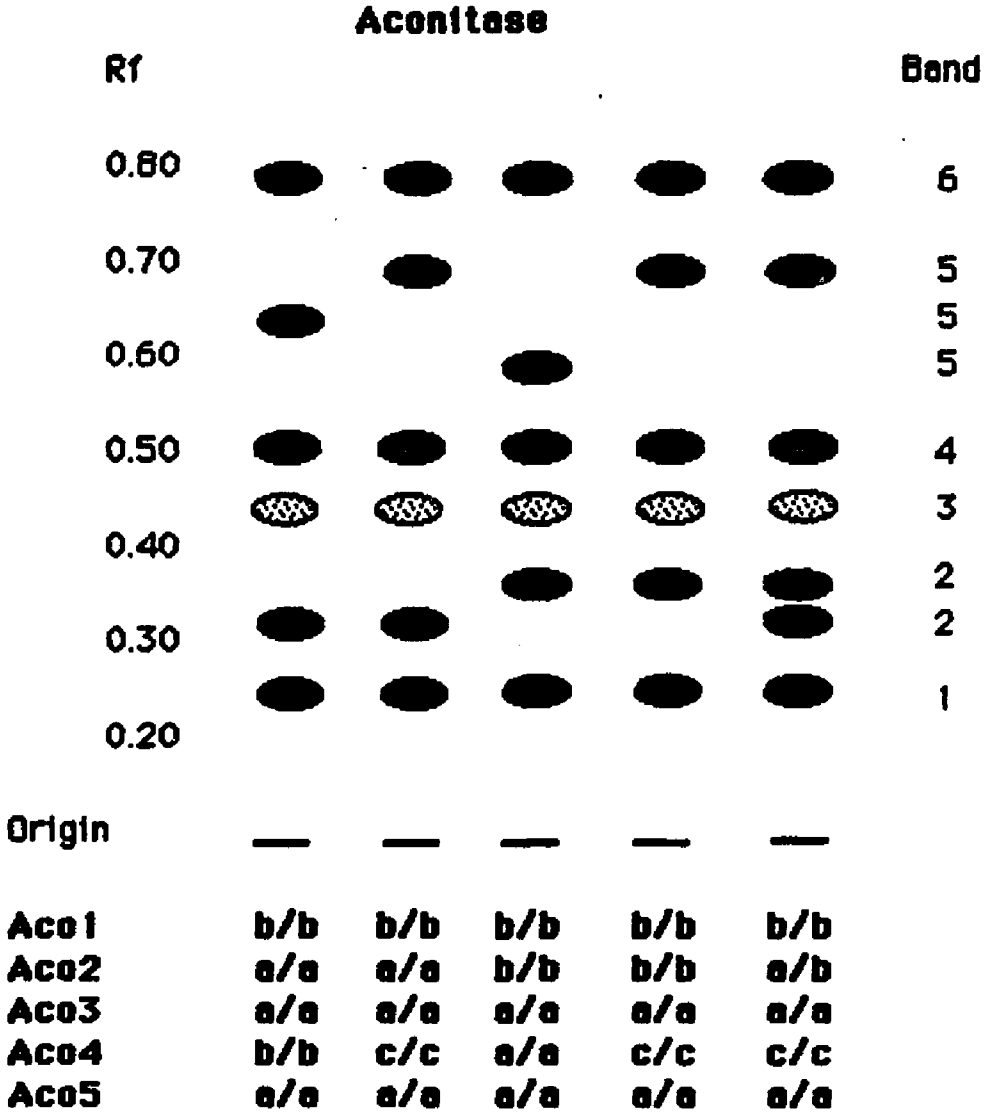



Figure 4. ACO banding patterns for G. soja seed tissue. Shading indicates band intensity.

(band 4, Figure 4) may be due to another, monomorphic, locus. More genetic data are needed to confirm this hypothesis.

Three mobility variants at band 5 were determined to be codominant alleles at a single locus, Aco4. The mobility of the alleles are: Aco4-a ($R_F$=0.57), -b ($R_F$=0.65) and -c ($R_F$= 0.69). Two mobility variants at band 6 were observed and were determined to be due to codominant alleles at a single locus, Aco-5: Aco5-a ($R_F$= 0.74) and -b ($R_F$=0.79). A null activity mutant was detected for band 6 in G. soja and it was hypothesized that the allele responsible for the variant is recessive to Aco5-a and -b (Chiang, 1985). Aco5-a is the most common allele in G. max and G. soja. Aco5-b was found in only one southern Japanese accession of G. soja (Chiang, 1985).

Four homozygous electromorphs for aconitase were detected in the present study (Figure 4).

3. Alcohol dehydrogenase (ADH). ADH isozymes mediate the reversible formation of ethanol and NAD from acetaldehyde and NADH. Hybridization and immunoelectrophoretic evidence demonstrated that ADH molecules are dimers in soybean (Beremand, 1975; Gorman and Kiang, 1978). ADH isozymes have been shown to function in the anaerobic stress response in some plants (Brown and Marshall, 1974).

Three homozygous ADH zymograms have been observed in G.

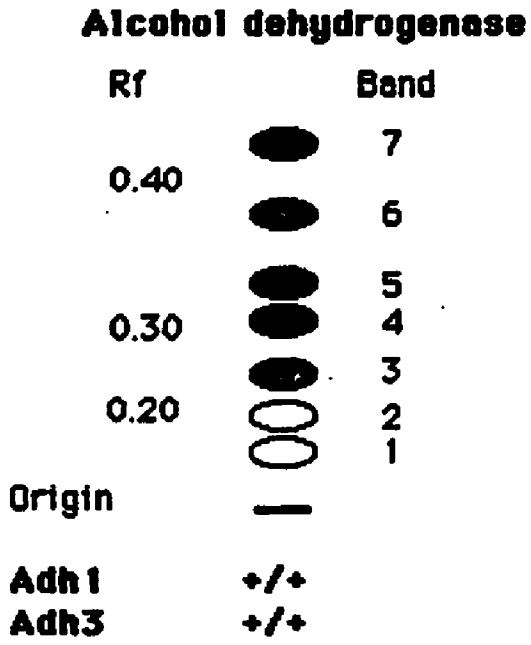



Figure 5. ADH banding pattern for G. soja seed tissue. Shading indicates band intensity. Open circles represent bands rarely observed.

max and G. soja (Gorman and Kiang, 1978; Gorman, 1983; Chiang, 1985). Gorman and Kiang (1978) proposed that bands 1, 3 and 5 are related in a fixed homo-heterodimer relationship between two loci, Adh1 and Adh2. That is, bands 1 ($R_F$ = 0.28) and 5 ($R_F$ = 0.40) are the homodimers of Adh1 and Adh2, respectively, and band 3 ($R_F$ = 0.34) is the heterodimer formed as the result of interaction of the monomers of Adh1 and Adh2 (Figure 5). No variants have been uncovered for the Adh2 locus and it was assumed to be homozygous and monomorphic in all individuals examined in the present study. A recessive null mutant (adh1) for the Adh1 locus was uncovered by Chiang (1985). When adh1 is expressed, bands 1 and 3 are not observed. Band 2 of the ADH zymogram is controlled by Adh3 and has a recessive null allele (adh3) and a dominant functional (Adh3) allele.

Only one electromorph was observed for G. soja seed in the present study (Figure 5). Five bands were observed consistently, although two additional weakly staining bands occasionally were observed near the origin.

4. Beta-amylase (AM). Beta-amylases hydrolyze starch by attacking alpha-1,4-glucosidic bonds in the linear chain amylose and in the branched amylopectin (Lehninger, 1982). Detection of a null activity mutant at the locus coding for beta-amylase suggests that these enzymes are not essential for normal starch metabolism in soybeans (Kiang, 1981).

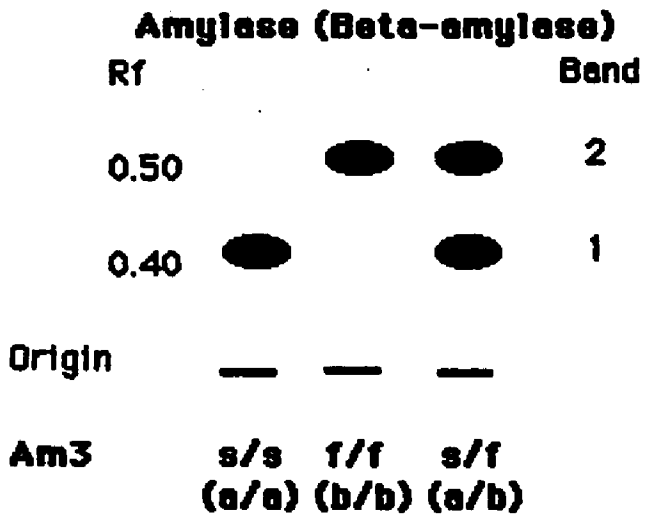



Figure 6. AM banding patterns for G. soja seed tissue.

Soybean amylases are monomeric (Morita et al., 1976).

Two single band mobility variants and two activity variants were detected for beta-amylase activity in G. max (Gorman and Kiang, 1977; Kiang, 1981). The two mobility variants were shown to be codominant alleles (Am3-s, $R_F$=0.41 and Am3-f, $R_F$=0.51) at one nuclear locus (Am3). These variants also are denoted Am3-a and Am3-b, respectively. One activity variant (Am3-sw) resulted in a slow, weakly staining band which subsequently was shown to be allelic and recessive to Am3-s and Am3-f, but dominant to a second, null, variant (Am3-n1) (Gorman and Kiang, 1978; Kiang, 1981).

Two mobility variants were observed in the present study of G. soja (Figure 6).

5. Diaphorase (DIA). Diaphorases (Lipoamide dehydrogenases) are a ubiquitous class of flavin-bound enzymes that cataylze the reduction of various dyes which act as hydrogen acceptors from dihydrodiphosphopyridine nucleotides (NADH) or dihydrotriphosphopyridine nucleotides (NADPH) (Dixon and Webb, 1979). The function of diaphorases in plants is unknown. To visualize DIA activity in vitro, the synthetic substrate, 2,6-dichlorophenol indophenol (DCIP), was employed as the electron acceptor in the present study.

The DIA zymogram patterns in wild and cultivated soybeans are complex, with as many as 12 anodal bands observed for

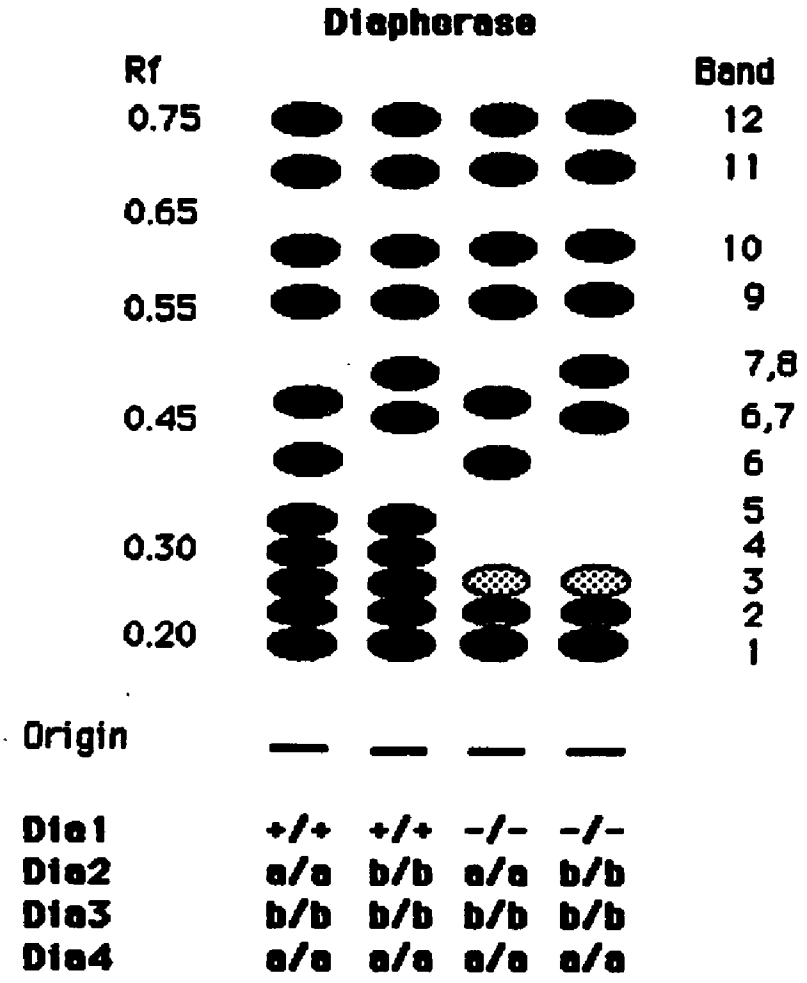



Figure 7. DIA banding patterns for G. soja seed tissue. Shading indicates band intensity. '+' and '-' indicate the dominant and recessive alleles, respectively.

dry seed cotyledon tissue. These bands were determined to be the products of five to seven loci (Gorman et al., 1983).

A cluster of five bands close to the origin are mitochondrial-associated and form a fixed homo-heterodimer complex, produced by two interacting loci (Gorman, 1983). Kiang and Gorman (1983) described the two alleles (Dia1 and dia1) at the Dia1 locus as 'incompletely dominant' since individuals homozygous for dia1 retained weak enzyme activity. No variants have been uncovered for the second locus hypothesized to be involved in the homo-heterodimer complex. For purposes of data analysis in the present study, this hypothetical locus was designated, Dia5, and was assumed to be homozygous and monomorphic in all individuals examined.

Dia2 has two mobility and one activity variants (Gorman and Kiang, 1983; Chiang, 1985). The mobility variants are codominant alleles which each produce a two band zymogram pattern: Dia2-a ($R_F$s=0.41 and 0.47) and Dia2-b ($R_F$s=0.46 and 0.51). A null activity variant (dia2) was detected in G. soja accessions from South Korea and was determined to be recessive to the other two alleles (Chiang, 1985).

Gorman et al. (1983) reported that Dia3 had a dominant and a recessive allele. Chiang (1985) detected a mobility variant at this locus in G. soja. She hypothesized that Dia3-a ($R_F$=0.56) is codominant to Dia3-b (formerly Dia3; $R_F$= 0.67) and that the activity variant is recessive to both of

these alleles. More genetic data are needed to confirm this hypothesis.

Genetic analysis of variants at the Dia4 locus revealed two codominant alleles were responsible for the observed zymogram patterns in G. soja. Dia4-a produced two bands, $R_F$s =0.68 and 0.75. Dia4-b produced one band, $R_F$=0.72. The reasons for the one band versus the two band zymograms are not yet understood.

Four homozygous electromorphs were observed in the present study (Figure 7).

6. Endopeptidase (ENP). Endopeptidases (pepsin, trypsin and chymotrypsin) are proteolytic enzymes which hydrolyze peptide bonds at interior positions within a polypeptide chain, making free alpha-amino positions available for exopeptidases. (Morris et al., 1985). High levels of endo- and exo- peptidases have been found in embyro and endosperm tissues of barley and wheat (Morris et al., 1985). It is hypothesized that these enzymes, in concert, break down storage proteins and provide an amino acid pool for seed germination, differentiation and growth (Morris et al., 1985). Trypsin is the most active on peptide bonds in which the carboxyl group is donated by an arginine or a lysine residue (Fairlet and Kilgour, 1966). Biochemical data indicate that these enzymes usually are monomeric in plants (Manchenko, 1988). A synthetic substrate, N-alpha-benzoyl

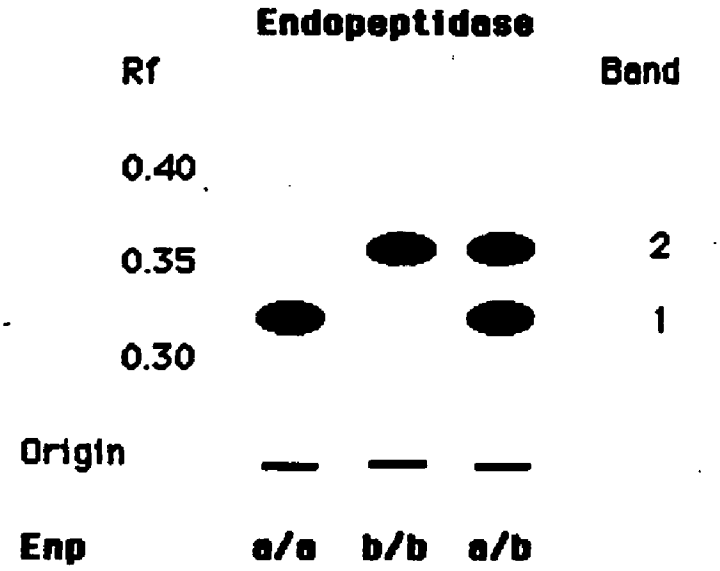



Figure 8. ENP banding patterns for G. soja seed tissue.

DL-arginine beta-naphthylamide (HCl) (BANA), was used *in vitro* to visualize ENP acitivity.

Three electrophoretic mobility variants have been observed in accessions of *G. max* and *G. soja* (Chiang, 1985; Doong, 1986). A single locus, *Enp*, with three codominant alleles is responsible for the observed zymogram patterns (Doong and Kiang, 1987b). *Enp-b* ($R_F$=0.36) is the most common allele in *G. soja* and *G. max* accessions, followed by *Enp-a* ($R_F$=0.33) (Doong and Kiang, 1987b). *Enp-c* ($R_F$=0.39) has been observed only in accessions of *G. soja* (Chiang, 1985).

*Enp-a* and *Enp-b* alleles were observed in the present study (Figure 8).

7. **Esterase (EST).** Carboxylesterases are a group of relatively non-specific enzymes capable of cleaving ester bonds. In plants, variant forms of esterases have been correlated with geographic elevation and latitude (Vickery and Hsu, 1984), plant growth habit (Darmency and Gasquez, 1983) and breakdown of seed storage proteins (Ferrer-Monge, 1974). Most plant esterases are monomeric (Manchenko, 1988).

Eight anodal bands and one cathodal band were observed for mature soybean seed in *G. max* (Bult and Kiang, 1989). The cathodal band showed two variants, a slow type ($R_F$=0.06) and a fast type ($R_F$=0.11). These were determined to be codominant alleles (*Est1-a* and *Est1-b*) at a single nuclear

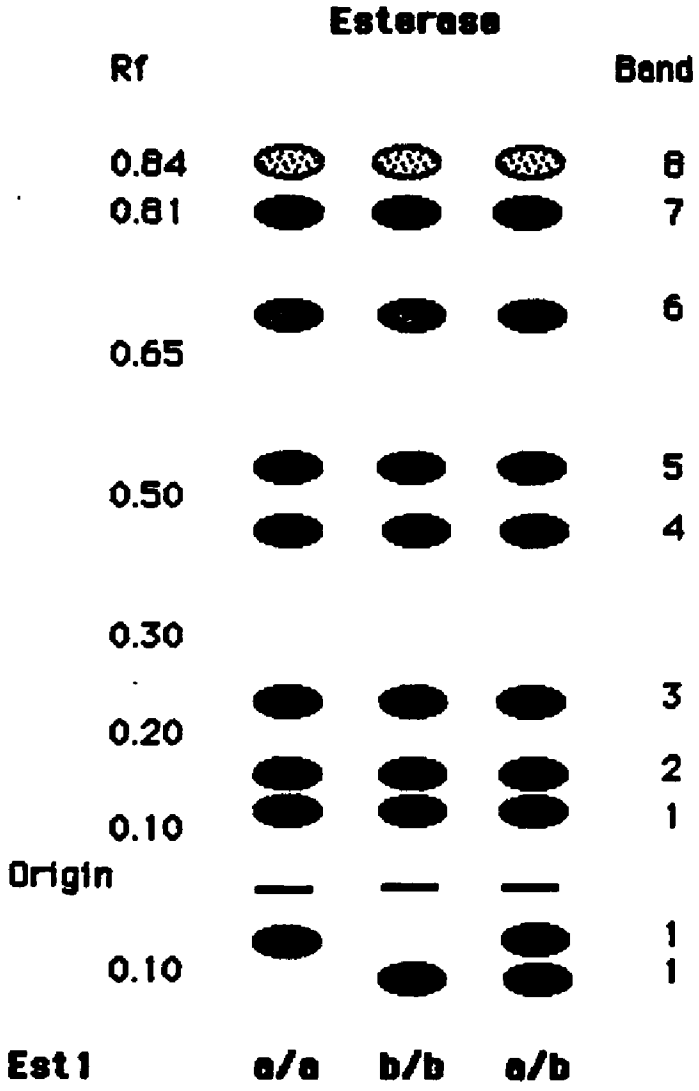



Figure 9. EST banding patterns for G. soja seed tissue. Shading indicates band intensity.

locus, Est1 (Bult and Kiang, 1989). Similar zymogram patterns were observed for G. soja seed in the present study (Figure 9).

8. Fluorescent esterase (FLE). Fluorescent arylesterases are involved in acetate biosynthesis and catalyze the following reaction:

$$\text{phenyl acetate} + H_2O \longrightarrow \text{phenyl} + \text{acetate}.$$

The subunit structure of fluorescent esterases has not been determined for most plants. Data for pinus species suggest FLE has a dimeric structure (Ryu, 1982).

Doong (1986) observed two homczygous zymogram types in G. max. Type 1 had five anodal bands. Type 2 lacked activity at band 1. Crosses between the two types indicated that band 1 was controlled by a single gene with Fle1 dominant to the recessive fle1 (Doong and Kiang, 1988). Because no variation was observed at the other bands, no genetic data are available.

A five-band pattern was observed for G. soja seed in the present study (Figure 10).

9. Glucose 6-phosphate dehydrogenase (GPD). GPD is a key enzyme in glucose metabolism in all known organisms (Lehninger, 1982). GPD mediates the first reaction in the pentose phosphate pathway, the dehydrogenation of glucose 6-phosphate.

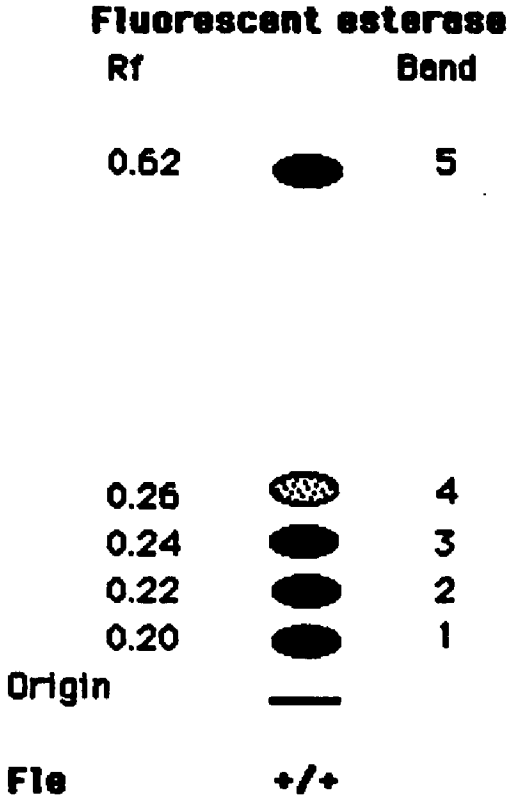



Figure 10. FLE banding pattern for G. soja seed tissue. Shading indicates band intensity. '+' indicates the dominant allele.

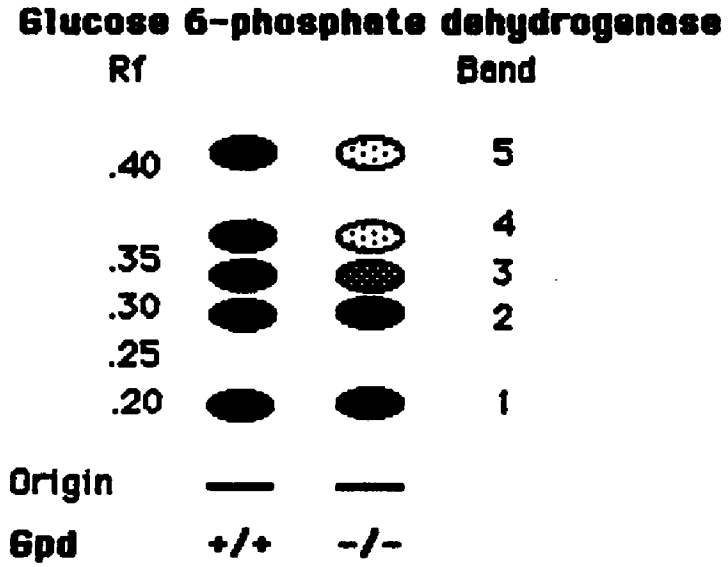



Figure 11. GPD banding patterns for G. soja seed tissue. Shading indicates band intensity. '+' and '-' indicate dominant and recessive alleles, respectively.

Two homozygous zymogram types have been detected among cultivated and wild soybeans which differ in intensities of bands 3, 4 and 5 (Figure 11). This variation is hypothesized to be controlled by a single nuclear locus with two variant alleles. Gpd produces the high intensity banding pattern and is dominant to gpd which produces the alternate zymogram (Gorman, 1983). Subcellular fractionation suggested that bands 3, 4 and 5 are cytosol-associated and that bands 1 and 2 are plastid-associated (Gorman, 1983).

The difficulty in interpreting zymogram patterns for GPD is that the enzyme has a regulatory role in a major biochemical pathway. Two possible genetic mechanisms for its expression were proposed by Kiang and Gorman (1983). The first possibility suggested by these workers is that one gene, Gpd, codes for bands 3, 4 and 5 and the electrophoretic variants are due to two alleles Gpd and gpd. A second possibility is that bands 3 and 5 are monomers of two separate gene loci and band 4 is a heterodimer. If the second possibility is true, then the Gpd locus (whose alleles differentiate the two observed electrophoretic types) may be a regulatory locus affecting the two GPD structural loci. For the purposes of this study, it was assumed that the first possibility is correct (see Gorman, 1983).

Two zymogram types were observed for G. soja seed in the present study (Figure 11).

10. Glutamate oxaloacetic transaminase (GOT). GOT (or aspartate aminotransferase, AAT) isozymes mediate the removal of amino groups from amino acids leading to the formation of keto acids for the Krebs cycle and gluconeogenesis (Lehninger, 1982; Kiang et al., 1987). It catalyzes the reversible reaction

aspartate + alpha-ketoglutarate $\rightleftharpoons$ oxaloacetate + glutamate.

These enzymes are specific to alpha-ketoglutarate or oxaloacetate as amino group acceptors, but are multispecific for amino group donors, with aspartate preferred (Whightman and Forrest, 1978). GOT has been reported as a dimer in plants (Manchenko, 1987).

Two zones of enzyme activity were observed in G. max and G. soja accessions (Gorman, 1983; Kiang et al., 1987). The zone closest to the origin consists of an invariable cluster of three bands ($R_F$'s = 0.26, 0.29, 0.32) which is associated with the cytosol (Gorman, 1983). Gorman (1983) hypothesized that two interacting loci are responsible for producing these bands, with the more darkly staining middle band (band 2) being a heterodimer. In the present study, these hypothetical loci were designated Got2 and Got3. It was assumed that both loci were homozygous and monomorphic in all individuals examined.

The fourth band, representing the plastid-associated enzyme, has three mobility variants. These variants were

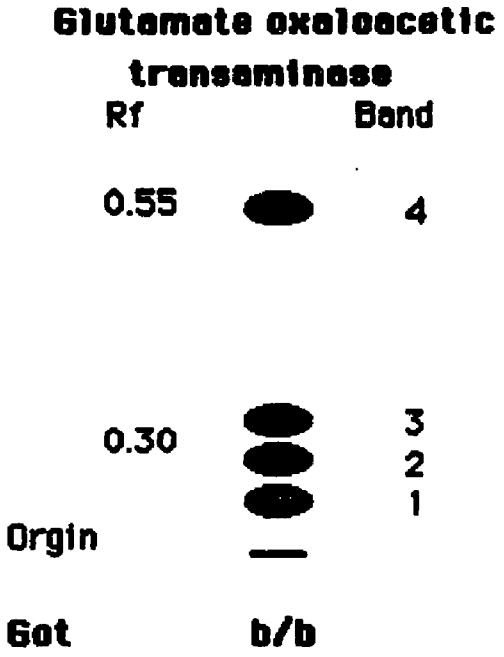



Figure 12. GOT banding pattern for G. soja seed tissue. Shading indicates band intensity.

found to be codominant alleles (Got-a, $R_F$=0.52; Got-b, $R_F$=0.54; Got-c, $R_F$=0.57) at the Got locus (Kiang et al., 1987).

Got-b is by far the most common allele, with Got-a and Got-c being very rare (Kiang et al., 1987). Only the Got-b allele was observed in the present study (Figure 12).

11. Isocitrate dehydrogenase (IDH). Two forms of IDH (NAD dependent and NADP dependent), catalyze the oxidation of isocitrate to alpha-ketoglutarate (Lehninger, 1982). The NAD-active IDH plays a major role in the citric acid cycle while the NADP-active form acts primarily to transfer reducing power between the mitochondria and the cytosol (Lehninger, 1982). Only the NADP-active form of IDH was examined in this study.

Zymogram patterns observed for dry cotyledon are due to two cytosol-associated loci (Idh1 and Idh2) and two mitochondrial-associated loci (Idh3 and Idh4) (Kiang and Gorman, 1985). Idh1 has two codominant alleles (Idh1-a, $R_F$=0.49 and Idh1-b, $R_F$=0.56). Idh2 has two codominant alleles (Idh2-a, $R_F$=0.56 and Idh2-b, $R_F$=0.63). A new variant at Idh2 was reported in a Japanese accession of G. soja and is inherited as a codominant allele at the Idh2 locus (Idh2-c, $R_F$=0.60) (Chiang, 1985). For Idh3, three codominant alleles were reported (Kiang and Gorman, 1985; Chiang, 1985): Idh3-a, $R_F$=0.31; Idh3-b, $R_F$=0.37; Idh3-c,

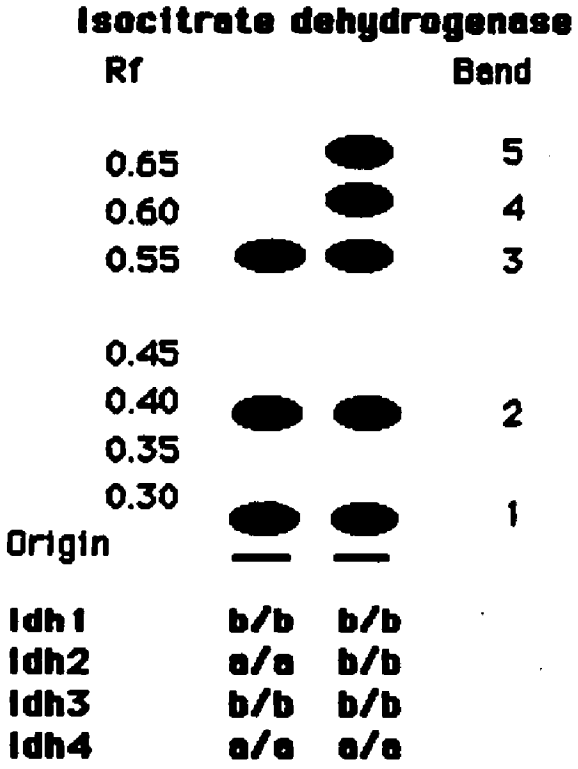



Figure 13. IDH banding patterns for G. soja seed tissue.

$R_F$=0.41).

Chiang (1985) reported two variants for the band closest to the origin and determined them to be codominant alleles at a single locus, Idh4. These variants are designated as Idh4-a ($R_F$=0.26) and Idh4-b ($R_F$=0.29). Chiang (1985) reported that the $F_2$ dihybrid progeny segregating for Idh2 and Idh4 showed no recombination, suggesting that the two loci are tightly linked or that the variants are the products of the same locus.

The two homozygous electromorphs observed for G. soja seed in the present study are shown in Figure 13.

12. Kunitz trypsin inhibitor (Ti). Kunitz trypsin inhibitor is an anti-nutritional seed protein present in most soybean lines (Hymowitz and Kaizuma, 1981). Three mobility and one activity variants have been observed in G. max. Genetic analysis revealed that the three mobility variants are allelic and co-dominant (Ti-a, $R_F$=0.79); Ti-b, $R_F$=0.75; Ti-c, $R_F$=0.83) (Hymowitz and Hadley, 1972). The null variant is inherited as a recessive allele (Orf and Hymowitz, 1979).

Ti-a and Ti-b alleles have been detected in G. soja (Chiang, 1985), with Ti-a being the most common allele. In the present study, only Ti-a was observed (Figure 14).

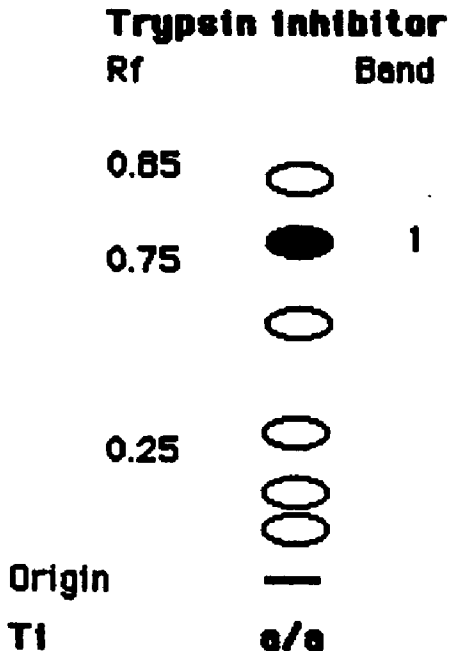



Figure 14. Ti banding pattern for G. soja seed tissue. Open circles represent non-Ti protein bands observed on the gel.

13. Leucine aminopeptidase (LAP). Leucine aminopeptidases are members of the exopeptidase enzyme group. They are relatively non-specific and hydrolyze peptide bonds adjacent to a free alpha-amino group of a peptide chain, providing a pool of free amino acids (Koehn et al., 1980). In vitro, the enzyme is visualized using the synthetic substrate L-leucine beta-naphthylamide (HCl) (Kiang et al., 1985). The lack of an intermediate band in $F_1$ and $F_2$ hybrid individuals suggests that this enzyme has a monomeric structure in soybeans (Kiang et al., 1985).

Two single-band variants were observed among G. max and G. soja accessions (Kiang et al., 1985) in dry cotyledon tissue. A single nuclear locus (Lap1) was found to control the two codominant alleles (Lap1-a, $R_F$=0.49 and Lap1-b, $R_F$=0.53) coding for the variants (Gorman, 1983).

A second LAP locus, Lap2, was described for 12- to 15-day-old seedlings in G. max and G. soja. This locus has two codominant mobility variants (Lap2-a, $R_F$=0.75 and Lap2-b, $R_F$=0.80) and one recessive activity variant (lap2) (Kiang and Chiang, 1987).

Only Lap1 was examined in this study. Both Lap1-a and Lap1-b variants were observed (Figure 15).

14. Malate dehydrogenase (MDH). Malate dehydrogenase mediates the final reaction in the citric acid cycle:

$$\text{L-malate} + \text{NAD} \longrightarrow \text{oxaloacetate} + \text{NADH}_2$$

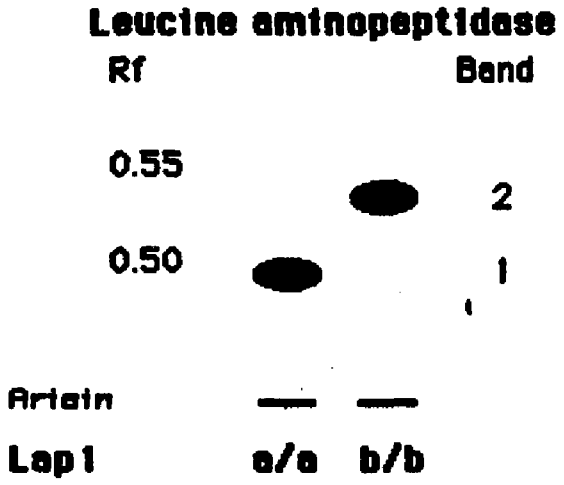
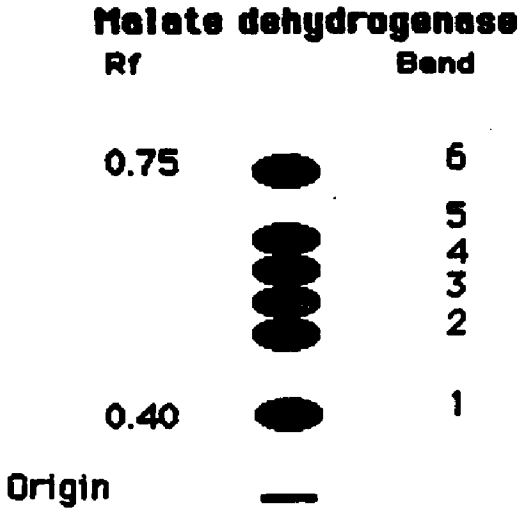



Figure 15. LAP banding patterns for G. soja seed tissue.




No genetic data are available.

Figure 16. MDH banding pattern for G. soja seed tissue.

(Lehninger, 1982). In all plant species examined, MDH was dimeric (Manchenko, 1988).

One MDH zymogram with six anodal bands was observed for G. soja in this study. Up to 10 bands have been reported among G. max and G. soja accessions (Gorman, 1983). Kiang and Gorman (1983) have hypothesized that three mitochondrial and three cytoplasmic associated bands are the products of two pairs of interacting loci. In the present study, the four hypothetical loci were assigned the designations, Mdh1, Mdh2, Mdh3 and Mdh4. It was assumed that these loci were homozygous and monomorphic in all individuals examined. No genetic data are available as no variants ever have been detected. A six band zymogram pattern was observed for MDH activity in the present study (Figure 16).

15. Mannose 6-phosphate isomerase (MPI). Mannose 6-phosphate isomerase mediates the isomerization of mannose 6-phosphate into fructose 6-phosphate, an intermediate in the central glycolytic pathway (Lehninger, 1982). Presence of both parental bands and no intermediate band in hybrid progeny suggests that this enzyme has a monomeric structure in soybeans (Gorman, 1983; Chiang, 1985).

All variants for MPI in soybean possess a single prominent band with a lighter co-migrating band which runs just anodal to the darker band. The genetic or epigenetic mechanisms responsible for these bands are not yet

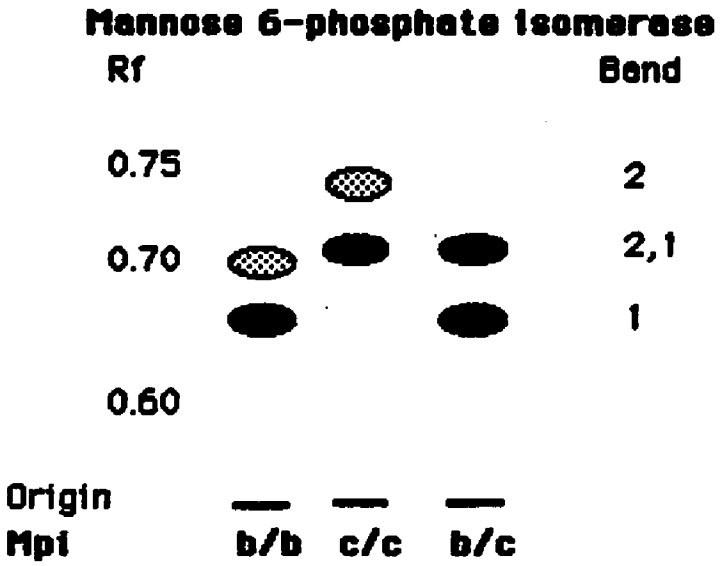



Figure 17. MPI banding patterns for G. soja seed tissue. Shading indicates band intensity.

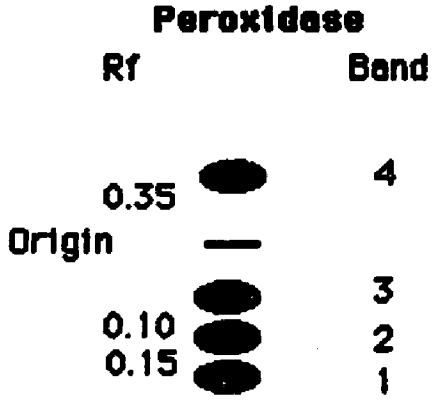
understood. No variants ever have been detected, and it is possible that the shadow band is due to some kind of post-transcriptional or post-translational modification or is an artifact of electrophoretic conditions.

Five homozygous zymograms have been visualized among G. soja and G. max accessions (Gorman et al., 1983). Four mobility variants have been determined to be allelic and codominant (Mpi-a, $R_F$=0.56 and 0.61; Mpi-b, $R_F$=0.56 and 0.70; Mpi-c, $R_F$=0.71 and 0.75 and Mpi-d, $R_F$=0.76 and 0.80)) (Kiang and Gorman, 1983; Chiang, 1985; Chiang and Kiang, 1988). A fifth activity variant, resulting in weak or null bands have been detected in G. max and is recessive to the other alleles at the Mpi locus (Kiang and Gorman, 1983; Chiang and Kiang, 1988).

Two variants, Mpi-b and Mpi-c were observed in the present study (Figure 17).

16. Peroxidase (PER). Peroxidases are ubiquitous in plant tissues, and multiple enzyme forms have been detected in many plant species (Buttery and Buzzell, 1968; Buzzell and Buttery, 1969). Although their exact physiological role is unknown, peroxidases are thought to be involved, directly or indirectly, in the regulation of indolacetic acid, cell wall formation, regulation of membrane permeability, disease resistance and seed dormancy (Buttery and Buzzell, 1968).

Buttery and Buzzell (1968) detected low and high levels




No genetic data are available.

Figure 18. PER banding pattern for G. soja seed tissue. Shading indicates band intensity.

of peroxidase activity in the seed coats of several G. max cultivars. They determined that seed coat peroxidase activity in soybean seed coats is controlled by a single locus, with a dominant allele (Ep) producing a high activity variant and a recessive allele (ep) producing low activity (Buzzell and Buttery, 1969).

A four band zymogram pattern (one anodal and three cathodal bands) was observed for G. soja seed in the present study and no variants were observed (Figure 18). I hypothesize that at least three genes are responsible for the seed peroxidase zymogram (Per1, Per2, Per3) and that two of these loci interact to form a homo-heterodimer complex (bands 1, 2 and 3). In the present study, these loci were assumed to be homozygous and monomorphic in all individuals tested.

17. 6-Phosphogluconate dehydrogenase (PGD). PGD isozymes mediate the oxidation/decarboxylation of 6-phosphogluconate in the oxidative phosphogluconate pentose phosphate pathway (Lehninger, 1982).

Four bands representing PGD activity generally are observed for dry soybean cotyledon tissue, although a two band pattern has been observed for some G. soja accessions (Chiang, 1985). Three loci are responsible for the zymogram pattern (Chiang and Kiang, 1987). Three mobility variants were reported for band 1 and are the products of codominant

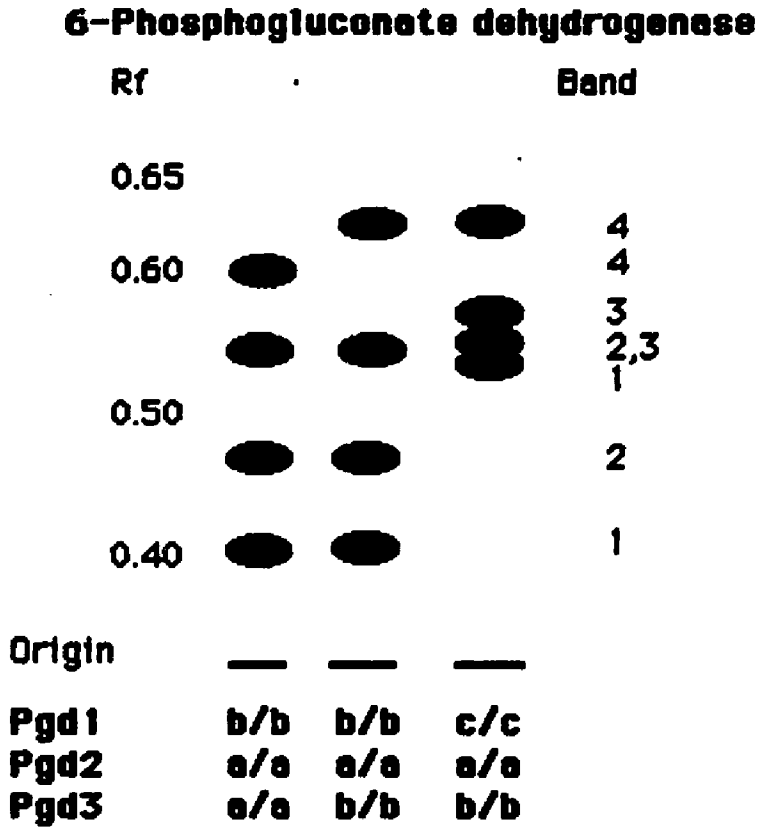



Figure 19. PGD banding patterns for G. soja seed tissue.

alleles at the Pgd1 locus (Pgd1-a, $R_F$=0.36; Pgd1-b, $R_F$=0.42; Pgd1-c, $R_F$=0.50). Pgd1-c only has been observed in G. soja. A null variant, pgd1, is hypothesized to exist based on observations by Chiang (1985) of a two band zymogram for G. soja seed accessions from Japan. More genetic data are needed to confirm this hypothesis.

Three mobility variants were observed for band 3 ($R_F$s=0.54, 0.62 and 0.68). Chiang (1985) demonstrated that these alleles are all codominant at a single nuclear locus, Pgd2. Band 2 always is midway between bands 1 and 3 and is hypothesized to be an interlocus heterodimer.

Band 4 has two mobility variants ($R_F$s=0.60 and 0.64). These variants were determined to be products of codominant alleles at a single nuclear locus, Pgd3. Pgd3-a has been detected only in a G. soja accession from Mishima, Japan (Chiang, 1985; Chiang and Kiang, 1987) and also was observed in the present study.

Three homozygous electromorphs were observed for PGD activity in the present study (Figure 19).

18. Phosphoglucose isomerase (PGI). Phosphoglucose isomerase (PGI) catalyzes the following reversible reaction (Lehninger, 1982):

$$\text{D-glucose 6-phosphate} \xrightleftharpoons{Mg_2^{++}} \text{D-fructose 6-phosphate}$$

The forward reaction functions in glycolysis, while the


Figure 20. PGI banding patterns for G. soja seed tissue. Shading indicates band intensity. '-' indicates the recessive allele.

reverse reaction functions in gluconeogenesis (Lehninger, 1982). PGI isozymes are dimers in many plant species, including soybean (Chiang, 1985; Manchenko, 1988).

As many as six anodal bands have been detected for PGI activity in G. max and G. soja. Two mobility and one activity variants were detected at the distal band (band 6). The two mobility variants were determined to be codominant alleles at a single nuclear locus, Pgi1 (Pgi1-a, $R_F$=0.58 and Pgi1-b, $R_F$=0.67). A null variant, pgi1, is recessive to these alleles. The products of the alleles coding for the mobility variants form an inter-locus heterodimer band (band 5) with the products of Pgi3. Pgi3 activity is responsible for band 3 and has two codominant alleles: Pgi3-a ($R_F$=0.42) and Pgi3-b ($R_F$=0.45).

A zone of enzyme activity closest to the origin displayed either a three-band or a single band patterns in G. soja and G. max. Chiang et al. (1987) hypothesized that two interacting loci produce a three band homo-heterodimer complex and that a null allele (pgi2) is responsible for the lack of band 1 and 2 in the single band variant (Figure 20). Band 4 may be the product of another Pgi locus, but this band stains inconsistently and has proven difficult to study genetically (Chiang, personal communication).

One four-band zymogram pattern was observed for G. soja seed in the present study (Figure 20).

19. Phosphoglucomutase (PGM). Like MPI, phosphoglucomutase is an important member of a "feeder" pathway that leads into the central glycolytic pathway (Lehninger, 1982). PGM catalyzes the reversible reaction,

glucose 1-phosphate glucose 6-phosphate.

PGM requires glucose 1,6 bisphosphate as a co-factor (Lehninger, 1982).

Two homozygous PGM zymograms have been observed for G. max and five in G. soja (Gorman et al., 1983; Chiang, 1985). At least three loci are hypothesized to be responsible for the observed banding patterns.

The zone of activity nearest to the origin has two single band mobility variants. These are due to a single nuclear locus, Pgm-1, with two codominant alleles (Pgm1-a, $R_F$=0.51 and Pgm1-b, $R_F$=0.54) (Kiang and Gorman, 1983). The banding pattern in the most distal bands (bands 2 and 3) is hypothesized to be controlled by two loci, Pgm2 and Pgm3. Pgm2 has three codominant alleles (Pgm2-a, $R_F$=0.64; Pgm2-b, $R_F$=0.67 and Pgm2-c, $R_F$=0.74). A weak to null activity variant detected at the third band in G. soja seed was demonstrated to be due to incompletely dominant alleles (Pgm3 and pgm3) at a third locus, Pgm3 (Chiang, 1985).

No variation was detected at Pgm2 or Pgm3 in the present study; both Pgm1 alleles were detected (Figure 21).

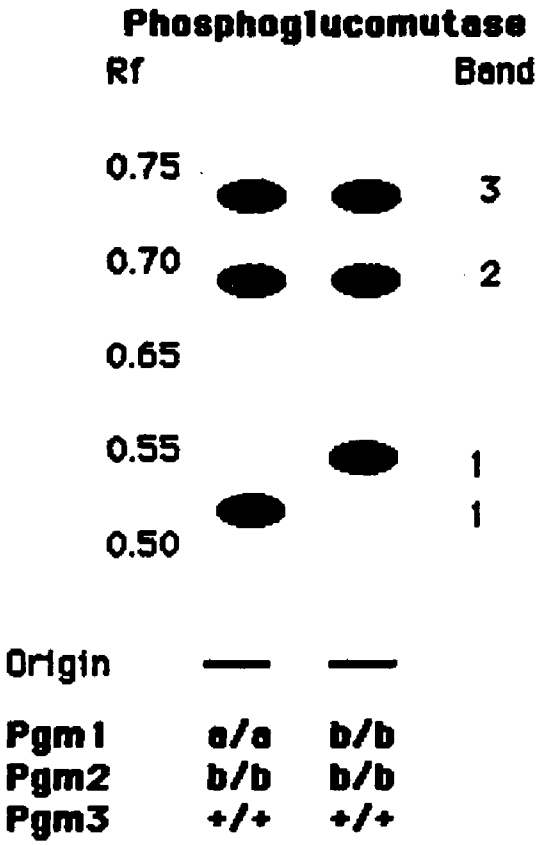


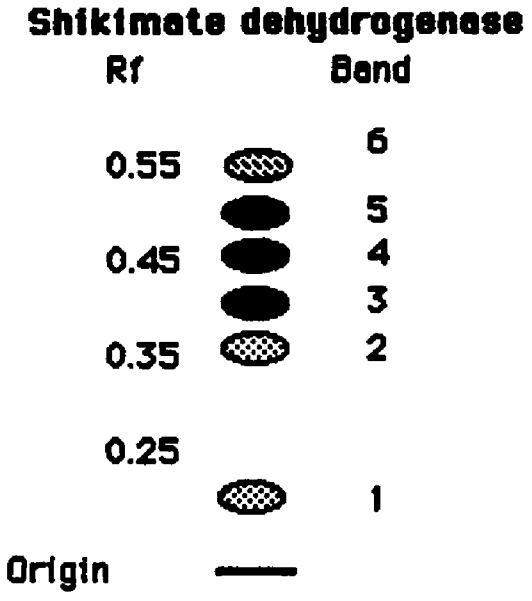

Figure 21. PGM banding patterns for G. soja seed tissue. '+' indicates the dominant allele.

20. Shikimate dehydrogenase (SKD). In plants, shikimate dehydrogenase acts in the biosynthetic pathways leading to prephenic and aromatic amino acids (Goodwin and Mercer, 1983). These aromatic ring containing molecules can, in turn, be used in alkaloid production as one plant defense against herbivory (Linhart et al., 1981). SKD catalyzes the reversible reaction,

shikimate + NADP $\rightleftharpoons$ 5-dehydroshikimate + NADPH.

Six anodal bands were evident in dry cotyledon of soybean, with bands 3,4 and 5 being the most intensely staining of the six. No mobility variants have been detected in G. max or G. soja. Chiang (1985) reported G. soja accessions with a consistently weak band 3 or 5, but no genetic data are available. Chiang (1985) proposed that bands 3 and 5 are the products of two homozygous loci with band 4 (which stains more intensely than do bands 3 or 5 ) as the heterodimer. In the present study, these hypothetical loci were assigned the designations Skd1 and Skd2 and were assumed to be homozygous and monomorphic in all individuals examined. A schematic of the zymogram observed for G. soja seed in the present study is shown in Figure 22.

21. Urease (EU). Ureases are thought to be involved in ureide metabolism in plants (Kerr et al., 1983). In soybeans, the activity of urease is correlated with that of arginase during germination, suggesting that urease is




No genetic data are available.

Figure 22. SKD banding pattern for G. soja seed tissue. Shading indicates band intensity.

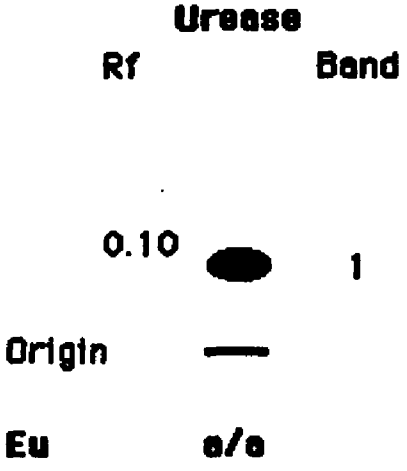



Figure 23. EU banding pattern for G. soja seed tissue.

involved in arginine metabolism possibly through the ornithine cycle (Kloth et al., 1987).

Polymers of a single subunit under different conditions form the electrophoretic mobility variants. Polacco and Havir (1979) determined that the embryo-specific urease can exist in at least two states of aggregation. The degree of aggregation of the subunit is controlled genetically.

Two mobility and one activity variants were detected among G. max and G. soja accessions (Buttery and Buzzell, 1971; Kloth and Hymowitz, 1985). Buttery and Buzzell (1971) proposed a genetic model of urease inheritance in which the fast form (Eu) was dominant to the slow form (eu). A re-examination of the inheritance at this locus by Kloth and Hymowitz (1985) demonstrated that the fast and slow soybean seed urease isozymes are associated with codominant alleles at a single locus. They suggest that Buttery and Buzzell consistently mis-identified heterozygotes in their analysis, because their water-based extraction procedure altered the heterozygote phenotype. Why the slow variant of seed urease synthesized in the heterozygote differs in stability or activity from the slow variant in homozygous form is not understood.

Only the slow type (Eu-a, $R_F$=0.09) was observed for G. soja seed in the present study (Figure 23).

**Genotype and allele frequencies based on zymogram patterns**

Genotype frequencies for the seven populations of wild soybean based on the observed zymogram patterns are summarized in Appendix IV. Allele frequencies for all loci (confirmed or hypothesized) detected in the present study are given in Appendix V. Table 3 summarizes all isozyme loci and alleles (confirmed or hypothesized) which have been detected to date in accessions or populations of G. soja.

B. Allele frequencies

Allele frequencies (Appendix V) were calculated from the genotype frequencies listed in Appendix IV. Fifteen of the 49 loci (30.6%) examined were polymorphic (99% criterion). A total of 31 alleles was observed at the variable loci (Table 4). Within any one population, no locus had more than two alleles. For most variable loci, one allele usually occurred at a much higher frequency (>0.60) than the other.

While the most frequent allele at a particular locus was the same for most of the populations, some loci were marked by changes in the predominant allele occurring over relatively short geographic distances (Table 4). For example, Am3-b was fixed in population 1 (Hatsunedai) but was less frequent or not found in the other populations. Dia1-b was the most frequent allele in population 3 (Nishi-asahiagaoka) and fixed in 4 (Yata), but was found at much lower frequencies in the other populations. Dia2-b and Lap-

Table 3. A summary of the isozyme loci and their alleles observed for G. soja seeds (references in text). Alleles marked with asterisks were observed in the present study.

| Enzyme symbol | No. of loci* | No. variable loci | Alleles observed at the variable loci |
|---|---|---|---|
| AP | 1 | 1 | Ap-a*, Ap-b, Ap-c*, Ap-d |
| ACO | 5 | 5 | Aco1-a, Aco-b*, Aco1-n; Aco2-a*, Aco2-b*, Aco2-c; Aco3-a*, Aco3-b; Aco4-a*, Aco4-b*, Aco4-c*; Aco4-d; Aco5-a*, Aco5-b |
| ADH | 3 | 2 | Adh1*, adh1; Adh3*, adh3 |
| AM3 | 1 | 1 | Am3-a*, Am3-b* |
| DIA | 5 | 4 | Dia1*, dia1*; Dia2-a*, Dia2-b*; Dia3-a, Dia3-b*; Dia4-a*, Dia4-b |
| ENP | 1 | 1 | Enp-a*, Enp-b*, Enp-c |
| EST1 | 1 | 1 | Est1-a*, Est1-b* |
| FLE | 1 | 1 | Fle*, fle |
| GPD | 1 | 1 | Gpd*, gpd* |
| GOT | 3 | 1 | Got-a, Got-b*, Got-c |
| IDH | 4 | 4 | Idh1-a, Idh1-b*; Idh2-a*, Idh2-b*; Idh3-a, Idh3-b*; Idh4-a*, Idh4-b |
| TI | 1 | 1 | Ti-a*, Ti-b, Ti-c, ti |
| LAP1 | 1 | 1 | Lap1-a*, Lap1-b* |
| MDH | 4 | 0 | |
| MPI | 1 | 1 | Mpi-a, Mpi-b*, Mpi-c*, Mpi-d |
| PER | 3 | 0 | |

Table 3. continued

| Enzyme symbol | No. of loci[a] | No. variable loci | Alleles observed at the variable loci |
|---|---|---|---|
| PGD | 3 | 3 | Pgd1-a, Pgd1-b*, Pgd1-c*, pgd1; Pgd2-a*, Pgd2-b, Pgd2-c; Pgd3-a*, Pgd3-b* |
| PGI | 3 | 3 | Pgi1-a, Pgi1-b*; Pgi2-a, Pgi2-b, Pgi2-c, pgi2*; Pgi3-a, Pgi3-b* |
| PGM | 3 | 3 | Pgm1-a*, Pgm1-b*; Pgm2-a, Pgm2-b*; Pgm3*, pgm3 |
| SKD | 2 | 0 | |
| EU | 1 | 2 | Eu-a*, Eu-b |
| Total | 48 | 36 | |

[a] Confirmed and hypothesized loci included.

Table 4. Frequencies for 31 alleles at 15 variable loci among seven natural populations of G. soja. N is the number of plants from which seeds were collected.

| | | Population | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Locus | N | 1 (13) | 2 (36) | 3 (10) | 4 (30) | 5 (10) | 6 (9) | 7 (3) |
| Ap | | | | | | | | |
| a | | 0 | 0 | 0 | 0 | 0.100 | 0 | 0 |
| c | | 1.000 | 1.000 | 1.000 | 1.000 | 0.900 | 1.000 | 1.000 |
| Aco2 | | | | | | | | |
| a | | 0.769 | 0.958 | 1.000 | 1.000 | 0.850 | 0.889 | 1.000 |
| b | | 0.231 | 0.042 | 0 | 0 | 0.150 | 0.111 | 0 |
| Aco4 | | | | | | | | |
| a | | 0.231 | 0.028 | 0 | 0 | 0 | 0 | 0 |
| b | | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 |
| c | | 0.769 | 0.972 | 1.000 | 0 | 1.000 | 1.000 | 1.000 |
| Am3 | | | | | | | | |
| a (s) | | 0 | 0.778 | 1.000 | 1.000 | 0.800 | 0.611 | 1.000 |
| b (f) | | 1.000 | 0.222 | 0 | 0 | 0.200 | 0.389 | 0 |
| Dia1 | | | | | | | | |
| Dia1 | | 0.231 | 0.056 | 0.900 | 1.000 | 0.300 | 0.111 | 0.333 |
| dia1 | | 0.769 | 0.944 | 0.100 | 0 | 0.700 | 0.889 | 0.667 |
| Dia2 | | | | | | | | |
| a | | 0.231 | 0.667 | 1.000 | 1.000 | 0.800 | 0.556 | 1.000 |
| b | | 0.769 | 0.333 | 0 | 0 | 0.200 | 0.444 | 0 |
| Enp | | | | | | | | |
| a | | 0.769 | 0.972 | 0 | 0 | 0.700 | 0.889 | 0.667 |
| b | | 0.231 | 0.028 | 1.000 | 1.000 | 0.300 | 0.111 | 0.333 |
| Est1 | | | | | | | | |
| a | | 0.231 | 0.861 | 0 | 0 | 0.500 | 0.833 | 0.667 |
| b | | 0.769 | 0.139 | 1.000 | 1.000 | 0.500 | 0.167 | 0.333 |
| Gpd | | | | | | | | |
| Gpd | | 0.692 | 0.222 | 0.900 | 1.000 | 0.400 | 0.111 | 0.333 |
| gpd | | 0.308 | 0.778 | 0.100 | 0 | 0.600 | 0.889 | 0.667 |

Table 4. continued

| Locus | N | Population 1 (13) | 2 (36) | 3 (10) | 4 (30) | 5 (10) | 6 (9) | 7 (3) |
|---|---|---|---|---|---|---|---|---|
| Idh2 | | | | | | | | |
| a | | 0 | 0 | 0 | 1.000 | 0.100 | 0.111 | 0 |
| b | | 1.000 | 1.000 | 1.000 | 0 | 0.900 | 0.889 | 1.000 |
| Lap1 | | | | | | | | |
| a | | 0.231 | 0.800 | 1.000 | 1.000 | 0.800 | 0.667 | 1.000 |
| b | | 0.769 | 0.200 | 0 | 0 | 0.200 | 0.333 | 0 |
| Mpi | | | | | | | | |
| a | | 1.000 | 1.000 | 1.000 | 1.000 | 0.850 | 0.889 | 1.000 |
| b | | 0 | 0 | 0 | 0 | 0.150 | 0.111 | 0 |
| Pgd1 | | | | | | | | |
| b | | 0.769 | 1.000 | 1.000 | 1.000 | 0.800 | 0.667 | 1.000 |
| c | | 0.231 | 0 | 0 | 0 | 0.200 | 0.333 | 0 |
| Pgd3 | | | | | | | | |
| a | | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 |
| b | | 1.000 | 1.000 | 1.000 | 0 | 1.000 | 1.000 | 1.000 |
| Pgm1 | | | | | | | | |
| a | | 1.000 | 1.000 | 1.000 | 0 | 1.000 | 1.000 | 1.000 |
| b | | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 |

b were the predominant alleles in Hatsunedai, whereas Dia2-a and Lap-a were more common in all other populations. Enp-b was fixed in Nishi-ashiagaoka and Yata, but was the less common allele in all the other areas. Est1-a was more common in populations 2, 6 and 7, while Est1-b was more common in 1, 3 and 4. In population 5, the two Est1 alleles occurred at equal frequencies (0.50). Gpd was more common in populations 1, 3 and 4, while gpd was more common in 2, 5, 6 and 7.

The Yata population (population 4) clearly was different from all other populations in two respects. First, the plants from this area were monomorphic at all loci examined. Second, the alleles at several loci were either unique to the Yata population (i.e., Aco4-b, Pgd3-a, Pgm1-b) or were fixed in Yata, but occurred at frequencies of less than 0.50 in most of the other populations (e.g., Idh2-a and, to some extent, Est-b, Dia1-b and Enp-b).

C. Genetic diversity within populations

The mean values of polymorphism (P) using the 99% criterion, number of alleles per isozyme locus (A), mean expected heterozygosity ($H_{exp}$), and proportion of heterozygotes observed ($H_{obs}$) over all seven wild soybean populations were 0.140, 1.14, 0.046 and 0.002, respectively (Table 5).

The measures of genetic variation varied considerably

Table 5. Within-population genetic diversity estimates for seven natural populations of wild soybean. N is the number of plants from which seed was collected in each population.

| Population | Sample size (N) | Mean No. of alleles per locus (A) | Proportion polymorphic loci (P) 95% | 99% | Heterozygosity[a] $H_{EXP}$ | $H_{OBS}$ |
|---|---|---|---|---|---|---|
| Hatsunedai | 13 | 1.19 (0.06) | 0.188 | 0.188 | 0.071 (0.02) | 0 |
| Kanodanchi | 36 | 1.19 (0.06) | 0.125 | 0.188 | 0.042 (0.02) | 0.002 (0.001) |
| Nishi-asahigaoka | 10 | 1.04 (0.03) | 0.042 | 0.042 | 0.008 (0.01) | 0 |
| Yata | 30 | 1.00 (0.00) | 0 | 0 | 0.000 (0.00) | 0 |
| Asahigaoka-1 | 10 | 1.25 (0.06) | 0.250 | 0.259 | 0.087 (0.02) | 0.004 (0.003) |
| Asahigaoka-2 | 9 | 1.25 (0.06) | 0.229 | 0.229 | 0.073 (0.02) | 0.005 (0.003) |
| Kakitagawa | 3 | 1.08 (0.04) | 0.083 | 0.083 | 0.044 (0.02) | 0 |
| Mean | | 1.14 | 0.131 | 0.140 | 0.046 | 0.002 |

[a] $H_{EXP}$ = Nei's gene diversity (1973); $H_{OBS}$ = the proportion of individuals in a population that were heterozygous.

among the populations (Table 5). P ranged from 0 in Yata (a completely monomorphic population) to 0.25 in Asahigaoka-1. The mean number of alleles ranged from 1.0 in Yata to 1.25 in Asahigaoka-1 and Asahigaoka-2. Estimates of mean expected heterozygosity ($H_{EXP}$) ranged from 0 in Yata to 0.087 in Asahigaoka-1. Population 6 (Asahigaoka-2) had the highest proportion of heterozygotes observed ($H_{OBS}$ = 0.005). The discrepancies between the expected and observed heterozygosity values in all populations are most likely due to the highly self-pollinating nature of soybean flowers.

D. Allele frequency heterogeneity

Heterogeneity contingency Chi-square tests demonstrated significant differentiation among the seven wild soybean populations at all variable loci (Table 6). Pgd3, Aco4, Pgm1 and Idh2 contributed most to population differentiation as indicated by their high Chi-square values. These loci all had alleles unique to population 4 (Yata). Some of the observed heterogeneity in allele frequencies was due to the variation in sample sizes among populations. Thus, the biological significance of allele frequency heterogeneity per se is questionable. For example, the relatively large sample size of population 4 (Yata), rather than some ecologically- or biologically-based phenomenon, may have resulted in a disproportionately large contribution to the overall Chi-square values for Pgd3, Pgm1, Idh2 and Aco4.

Table 6. Allele frequency heterogeneity of 15 isozyme loci observed among seven natural populations of G. soja.

| Locus | Allele | Mean | Variance | DF | $X^2$ * |
|---|---|---|---|---|---|
| Ap | a | 0.021 | 0.038 | 6 | 20.38 |
| | c | 0.979 | 0.038 | | |
| Aco2 | a | 0.924 | 0.090 | 6 | 21.98 |
| | b | 0.076 | 0.090 | | |
| Aco4 | a | 0.037 | 0.086 | 12 | 252.51 |
| | b | 0.143 | 0.378 | | |
| | c | 0.820 | 0.371 | | |
| Am3 | a | 0.741 | 0.359 | 6 | 112.38 |
| | b | 0.259 | 0.359 | | |
| Dia1 | Dia1 | 0.419 | 0.377 | 6 | 151.01 |
| | dia1 | 0.581 | 0.377 | | |
| Dia2 | a | 0.751 | 0.249 | 6 | 72.18 |
| | b | 0.249 | 0.290 | | |
| Enp | a | 0.571 | 0.404 | 6 | 165.74 |
| | b | 0.429 | 0.404 | | |
| Est1 | a | 0.442 | 0.369 | 6 | 132.33 |
| | b | 0.558 | 0.369 | | |
| Gpd | Gpd | 0.523 | 0.344 | 6 | 102.51 |
| | gpd | 0.477 | 0.344 | | |
| Idh2 | a | 0.173 | 0.368 | 6 | 204.56 |
| | b | 0.827 | 0.368 | | |
| Lap1 | a | 0.785 | 0.277 | 6 | 76.15 |
| | b | 0.215 | 0.277 | | |
| Mpi | a | 0.963 | 0.065 | 6 | 25.42 |
| | b | 0.037 | 0.065 | | |
| Pgd1 | b | 0.891 | 0.142 | 6 | 45.33 |
| | c | 0.109 | 0.142 | | |
| Pgd3 | a | 0.143 | 0.378 | 6 | 222.00 |
| | b | 0.857 | 0.378 | | |
| Pgm1 | a | 0.857 | 0.378 | 6 | 222.00 |
| | b | 0.143 | 0.378 | | |

* All Chi-square values significant at the 1% level.

## E. Partitioning genetic variation within and among populations

The question of how genetic variation is partitioned within and among the seven populations was explored using Nei's genetic diversity measures (1973) (Table 7). The diversity measures were weighted by population size.

Total genetic diversity ($H_T$) for the species ranged from 0.499 for Gpd to 0.028 for Ap. The mean value over all fourteen variable loci was 0.306. Within-population genetic diversity ($H_S$) ranged from 0 for Pgm1 and Pgd3 to 0.254 for Dia2. The mean value for $H_S$ over all variable loci was 0.113. This value is very similar to the mean $H_S$ value of 0.128 for 39 autogamous plant species (Loveless and Hamrick, 1984) and 0.126 for 13 inbreed plant species examined by Brown (1979).

$G_{ST}$ measures the proportion of variation among populations relative to the total species' diversity (Loveless and Hamrick, 1984). The $G_{ST}$ values in this study ranged from 0.243 at Aco2 to 1.0 at Pgm1 and Pgd3. The overall value of $G_{ST}$ was 0.627 (calculated as the average $D_{ST}$ divided by the average $H_T$). Thus, over 60% of the total genetic variation resided among populations. This value is almost three times that observed by Chiang (1985) among four natural populations of wild soybean from Japan ($G_{ST}$ = 0.198), but is only slightly higher than the mean $G_{ST}$ value of 0.523 for 30 autogamous species examined by Loveless and Hamrick (1984).

Table 7. Partitioning of genetic variation within and among seven populations of wild soybean.

| Locus | Total diversity $H_T$ | Within population $H_S$ | Between population $D_{ST}$ | Gene differentiation $G_{ST}$ |
|---|---|---|---|---|
| Aco2 | 0.140 | 0.107 | 0.034 | 0.243 |
| Aco4 | 0.306 | 0.059 | 0.246 | 0.804 |
| Am3 | 0.386 | 0.179 | 0.204 | 0.528 |
| Ap | 0.028 | 0.016 | 0.011 | 0.393 |
| Dia1 | 0.487 | 0.158 | 0.329 | 0.676 |
| Dia2 | 0.374 | 0.254 | 0.120 | 0.321 |
| Enp | 0.490 | 0.125 | 0.365 | 0.745 |
| Est1 | 0.493 | 0.158 | 0.329 | 0.667 |
| Gpd | 0.499 | 0.249 | 0.250 | 0.501 |
| Idh2 | 0.286 | 0.032 | 0.254 | 0.888 |
| Lap1 | 0.346 | 0.213 | 0.134 | 0.387 |
| Mpi | 0.071 | 0.039 | 0.032 | 0.451 |
| Pgd1 | 0.194 | 0.109 | 0.084 | 0.433 |
| Pgd3 | 0.245 | 0 | 0.245 | 1.000 |
| Pgm1 | 0.245 | 0 | 0.245 | 1.000 |
| Mean | 0.306 | 0.113 | 0.192 | 0.627[a] |

[a] calculated as $D_{ST}/H_T$

Why is the $G_{ST}$ value in the present study so much higher than that observed by Chiang (1985)? The study populations examined by Chiang were collected along a 120 km stretch of the Kitakami River on the northern island of Honshu, Japan (1985). If seed flow occurred down the river it would tend to have a homogenizing effect of the genetic structure among populations, lowering the overall proportion of variation among populations relative to total genetic diversity.

In the present study, the average distance between collection sites was only 1.54 km. However, migration between populations by means of some natural vector was not likely. Therefore, the seven soybean populations examined here are likely more isolated genetically than those examined by Chiang. In addition, the population and sample sizes examined here were significantly smaller than those in Chiang's study (1985). Thus, founder effects and genetic drift may be playing a larger role in maintaining differences among the soybean populations examined here.

F. Genetic distance

Nei's (1972) genetic distance (D) and identity (I) measures were calculated for pairwise comparisons among the seven soybean populations (Table 8). The mean values for D and I were 0.063 and 0.941, repectively (Table 9). The mean values for genetic distance and identity among the four natural populations of G. soja examined by Chiang (1985)

Table 8. Matrix of Nei's (unbiased) genetic identity (above the diagonal) and distance (below the diagonal) coefficients among seven natural populations of wild soybean.

| Population | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | ***** | 0.959 | 0.927 | 0.841 | 0.969 | 0.970 | 0.946 |
| 2 | 0.042 | ***** | 0.935 | 0.844 | 0.994 | 0.998 | 0.995 |
| 3 | 0.076 | 0.067 | ***** | 0.916 | 0.969 | 0.930 | 0.971 |
| 4 | 0.173 | 0.170 | 0.088 | ***** | 0.881 | 0.842 | 0.881 |
| 5 | 0.032 | 0.006 | 0.032 | 0.127 | ***** | 0.995 | 1.000 |
| 6 | 0.031 | 0.002 | 0.072 | 0.172 | 0.005 | ***** | 0.989 |
| 7 | 0.055 | 0.005 | 0.029 | 0.127 | 0.000 | 0.011 | ***** |

Table 9. Mean genetic distance ($\bar{D}$) and identity ($\bar{I}$) among seven populations of G. soja from Mishima, Japan.

| | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Average |
|---|---|---|---|---|---|---|---|---|
| $\bar{D}$ | 0.068 | 0.049 | 0.061 | 0.143 | 0.034 | 0.049 | 0.038 | 0.063 |
| $\bar{I}$ | 0.935 | 0.954 | 0.941 | 0.868 | 0.968 | 0.954 | 0.964 | 0.941 |

Table 10. Spearman rank correlations between genetic distance and estimated geographic distance among seven natural populations of wild soybean.

| Population | Mean genetic distance (D) | Mean geographic distance[a] (km) | Correlation coefficient[b] |
|---|---|---|---|
| 1. Hatsunedai | 0.068 | 1.08 | 0.377 |
| 2. Kanodanchi | 0.049 | 1.25 | 0.232 |
| 3. Nishi-asahigaoka | 0.061 | 1.13 | -0.169 |
| 4. Yata | 0.143 | 1.54 | -0.134 |
| 5. Asahigaoka-1 | 0.034 | 1.08 | -0.866 |
| 6. Asahigaoka-2 | 0.049 | 1.25 | 0.377 |
| 7. Kakitagawa | 0.038 | 3.42 | -0.530 |
| Overall Mean | 0.063 | 1.54 | 0.037 |

[a] Distances between populations estimated from map of soybean seed collection sites provided by H.I. Oka.
[b] All correlation coefficients are not significantly different from zero at the 5% level.

were 0.444 and 0.958, respectively. In the present study, populations 5 (Asahigaoka-1) and 7 (Kakitagawa), two of the most geographically distant populations (over 3 km), were genetically identical based on allele frequencies. The maximum divergence observed was between populations 1 (Hatsunedai) and 4 (Yata) (D=0.173, I=0.841).

Geographic distances between populations were estimated from a map of sample areas provided by the collectors (Drs. Oka and Morishima). The spatial distance from one population to all others was averaged and compared with Nei's genetic distance coefficients. The genetic distance measures did not correlate significantly with geographic distance, suggesting that isolation by distance is not a large component of genetic differentiation among these populations (Table 10). Similar results have been observed for other predominantly selfing plant species, including Avena barbata (slender wild oat, Kahler et al., 1980) and Hordeum spontaneum (wild barley, Nevo et al., 1979).

G. Canonical discriminant analysis

Relationships among the seven populations were analyzed further using canonical discriminant analysis. The between canonical functions (which indicate those variables most useful for discriminating among populations) and the standardized canonical coefficients (which are similar to regression coefficients) for the electrophoretic data are

given in Appendix VI.

Four of the six canonical variates extracted from the 31-member isozyme allele data matrix were significant at the 1% level as indicated by Wilk's lambda. The eigenvalue for the first canonical variate was incalculable because the within-population variation for this variate was the value, zero. An eigenvalue is analagous to a ratio of between-class sums of squares to within-class sums of squares (Norusis, 1985). In general, a canonical variate with a large eigenvalue is one that accounts for a significant proportion of the total variance in the model attributable to between-class differences (Norusis, 1985). For the first canonical variate in the present study, all the variation was due to the between-class component. Thus, although an eigenvalue could not be computed, the variate was very important in differentiating among populations. The loci which contributed most to this function (as indicated by between canonical structure) were Aco4, Idh2, Pgd3, Pgm1, Enp, Est and Gpd. These loci had alleles which were unique to population 4 (Yata) or which were fixed in Yata but occurred at frequencies below 50% in most of the other wild soybean populations.

The other three statistically significant canonical variates accounted for 55.7%, 37.0% and 4.8% of the total variation in the model, respectively. The loci which contributed most to the second canonical variate were Aco2,

Aco4, Am3, Lap, Dia2 and Enp. The loci which contributed most to the third variate were Aco4, Est and Gpd. Two loci, Ap and Mpi contributed most to the fourth canonical variate.

A scattergram of the class means on the first two canonical variates is shown in Figure 24. The first canonical variate (x-axis) separated population 4 (Yata) from all other populations of wild soybean based largely on the presence of the several alleles which were unique to Yata. The second canonical variate (y-axis) separated the remaining six populations based on differences in the frequencies of alleles among populations. As population 4 was completely monomorphic, its y-axis intercept was the value, zero.

Three or four population groupings are evident from the scattergram (Figure 24). Population 1 (Hatsunedai) was grouped by itself, as were populations 3 (Nishi-asahigaoka) and 4 (Yata). Populations 2 (Kanodanchi), 5 (Asahigaoka-1) and 6 (Asahigaoka-2) were grouped tightly together. Population 7 (Kakitagawa) was close to the cluster of populations 2, 5 and 6, but could also be considered as a separate group.

## H. Patterns of genetic variation in relation to environmental variables

Topography, shading and percent relative area occupied by wild soybean plants were estimated based on collectors' observations (See Appendices I and II). Relationships among

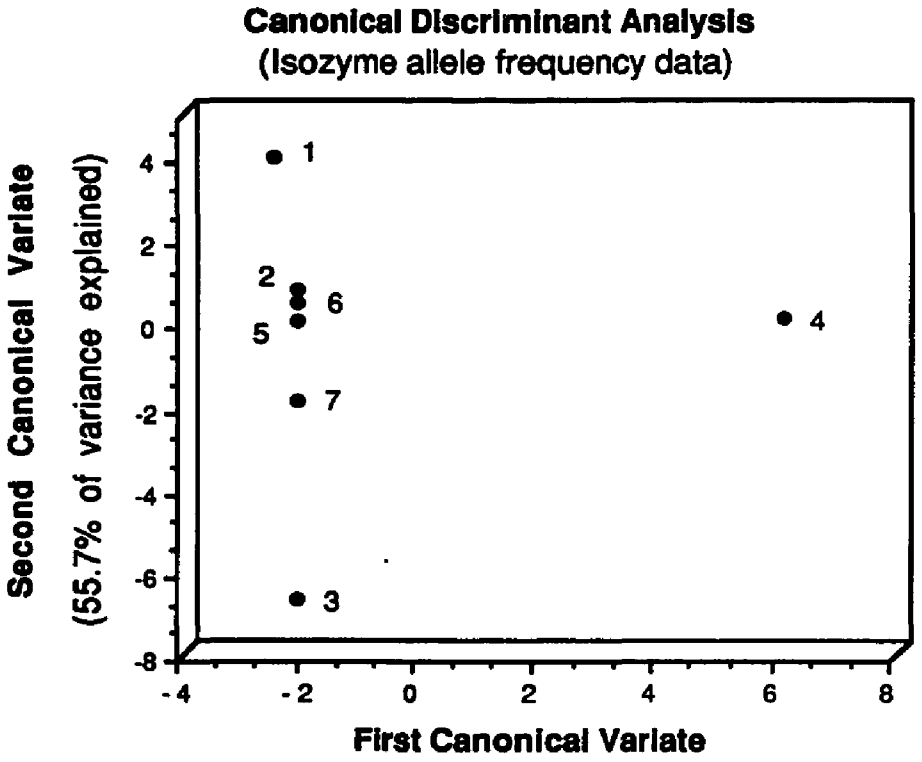



Figure 24. Scattergram of population means on the first two canonical variates. Based on a canonical discriminant analysis of a 31-member isozyme allele frequency data matrix. Populations: 1=Hatsunedai; 2=Kanodanchi; 3=Nishi-asahigaoka; 4=Yata; 5=Asahigaoka-1; 6=Asahigaoka-2; 7=Kakitagawa.

these environmental variables and Nei's genetic distance (D), number of alleles per locus (A), polymorphism ($P_{95}$ and $P_{99}$), and expected and observed mean heterozygosity ($H_{EXP}$ and $H_{OBS}$) were explored for the seven soybean populations using Spearman rank correlation analysis. Significant negative correlations were observed between percentage of the collection site occupied by wild soybean (relative to other, co-habitating plant species; COVER) and the number of alleles per locus (A) ($r_s$=-0.83), polymorphism ($r_s$=-0.77 and -0.83 for the 95 (P95) and 99% (P99) levels, respectively), and observed mean heterozygosity ($H_{OBS}$) ($r_s$=-0.88) (Table 11).

Table 11. Spearman rank correlation coefficients among environmental variables and genetic diversity estimates.

| | Environmental variables[a] | | | | Genetic diversity estimates[b] | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | TOP | SHADE | DIST | COVER | D | A | P95 | P99 | $H_{EXP}$ | $H_{OBS}$ |
| TOP | ____ | -0.71 | -0.33 | 0 | -0.59 | 0.10 | 0.10 | 0.10 | 0.29 | 0.12 |
| SHADE | | ____ | 0 | -0.08 | 0.40 | -0.16 | -0.16 | -0.16 | -0.32 | -0.26 |
| DIST | | | ____ | 0.82 | 0.77 | -0.82 | -0.77 | -0.82 | -0.77 | -0.33 |
| COVER | | | | ____ | 0.46 | -0.83*[c] | -0.77* | -0.83* | -0.60 | -0.88** |
| D | | | | | ____ | -0.66 | -0.61 | -0.66 | -0.67 | -0.56 |
| A | | | | | | ____ | 0.99** | 1.00** | 0.94** | 0.83* |
| P95 | | | | | | | ____ | 0.99** | 0.96** | 0.77* |
| P99 | | | | | | | | ____ | 0.94** | 0.83* |
| $H_{EXP}$ | | | | | | | | | ____ | 0.67 |

[a] TOP=topography; SHADE=shading; DIST=site disturbance; COVER-percent site occupied by wild soybean. See Appendices I and II for details.

[b] D=Nei's genetic distance; A=number of alleles per locus; P95 and P99=polymorphism at the 95% and 99% levels, respectively; $H_{EXP}$=gene diversity; $H_{OBS}$=proportion heterozygotes observed.

[c] * and **: correlation coefficient is significantly different from zero at the 5% and 1% levels, respectively.

# DISCUSSION

## Summary

The isozyme data revealed that the seven populations of Glycine soja examined here maintain most of their genetic variation among populations and that populations are well differentiated genetically. The levels and patterns of genetic variation within and among the populations of G. soja are consistent with the highly self-pollinating breeding system of this species and the severely restricted gene flow among populations (Loveless and Hamrick, 1984).

## Within-population genetic diversity

The within-population estimates of genetic variation for the seven populations of G. soja examined in this study are lower than those obtained from previous studies of natural populations of G. soja (Chiang, 1985) and accessions of G. soja from diverse geographical locations (Kiang et al., 1987). Chiang (1985) reported P, A and $H_{EXP}$ vlaues of .381, 1.55 and 0.114 for G. soja seed from four natural populations collected along the Kitakami River in the northern region of Honshu Island, Japan. Kiang and coworkers (1987) reported overall genetic diversity values of 0.622, 2.00 and 0.15, respectively, for wild soybean accessions collected from broad geographic regions.

These results are not suprising given the restricted sampling areas, small sample sizes, and relatively small sizes of the soybean populations examined in this study. Three to thirty-six individuals per population from relatively small populations were collected, restricting the proportion of the total genetic variation which could be represented in any one sample. The average distance between collection sites was only 1.54 km. By contrast, Chiang's (1985) sampling area was over 120 km, and the accessions examined by Kiang et al. (1987) represent worldwide collections of wild soybeans. Sample sizes in both instances were much larger than in the present study. Thus, these researchers had a greater chance of including more genetic variation (including rare alleles) in their soybean collections.

The within-population levels of genetic variation in the present study of G. soja also were lower than levels observed for other annual, selfing plants (Hamrick et al., 1979). The mean value for polymorphism (99%) in the present study was 0.140, slightly lower than the mean value of 0.179 for 33 selfing plant species examined by Hamrick et al. (1979). The mean number of alleles per locus in the present G. soja study was 1.14 compared to an average of 1.27 for other selfing species (Hamrick et al., 1979). Finally, the expected mean heterozygosity in the present study was 0.046 which is somewhat lower than the average value of 0.058

observed by Hamrick and coworkers (1979). The differences between the values in this study and those obtained by Hamrick et al. may be due to the relatively small sample population and sample sizes used in the present study and the number and types of loci included in the analysis.

The relatively low levels of intrapopulation genetic diversity and skewed nature of the allele frequencies G. soja is consistent with a condition of restricted gene flow among the populations which is, in turn, associated with the highly self-pollinating nature of the soybean flowers and the annual life cycle of this species (Hamrick et al., 1979; Loveless and Hamrick, 1984; Slatkin, 1987).

## Population differentiation

Like many annual, selfing plant species, G. soja maintains the majority of its genetic variation among rather than within populations. Wright (1969) stated that significant genetic differentiation was indicated when $F_{ST}$ (a measure similar to $G_{ST}$) was greater than 0.33. The average $G_{ST}$) in this this study was 0.627. Thus, the soybean populations can be considered highly differentiated.

Geographic distance proved to be an unreliable predictor of genetic similarity among the wild soybean populations. Isolation by distance, therefore, does not appear to be a major factor affecting population differentiation. Instead, the patterns of differentiation are characteristic of

ecotype formation in response to environmental heterogeneity of the type observed for Avena barbata (Clegg and Allard, 1972; Kahler et al., 1980; Allard, 1988), Hordeum spontaneum (Nevo et al., 1979), and Triticum dicoccoides (Nevo et al., 1982).

In the present study, there are four major lines of electrophoretic evidence which suggest ecotype formation in the wild soybean populations. Although this evidence is largely indirect, it is consistent with the patterns observed in other selfing, annual plant species for which ecotype formation has been demonstrated conclusively (Allard, 1988; Nevo, 1987). First, changes in the most frequent allele among some populations occurred over relatively short geographic distances. Second, one soybean population, Yata, possessed alleles not present in any of the other populations. Third, simple correlations between Nei's genetic distance and geographic distance were non-significant in all instances. Other processes, including founder effects, genetic drift and gene hitch-hiking also may be playing a role in conditioning genetic structure in the wild soybean populations (Hedrick, 1986).

### Genetic variation and environmental variables

Correlations between gene frequencies and environmental factors have been observed over macro- and micro- geographic scales in natural and experimental populations of plants

(Nevo, 1978; Allard, 1988). Even in the face of extensive gene flow, significant differentiation among populations and ecotype formation have been observed to occur over distances of less than one square meter in some populations of plants (Bradshaw, 1972; Kahler et al., 1980; Jain and Bradshaw, 1966; Antonovics and Bradshaw, 1970; Kiang, 1982). A local population, therefore, may be a mosaic of subpopulations or neighborhoods (Brown, 1979). This pattern is characteristic of natural populations of inbreeding plants (Allard et al., 1968).

Detailed information on environmental variables is not available for the wild soybean populations examined in this study, so it is not possible to explore the associations between genetic variation and environmental heterogeneity, except in the broadest possible terms. One interesting association indicated by simple correlation coefficients was the negative association between amount of genetic variation and the percentage of sampling area occupied by wild soybeans relative to other, co-habitating plant species (COVER). In Hatsunedai, Kanodanchi, Asahigaoka-1 and Asahigaoka-2, the percentage of soybean cover was low, from 5% to 30% (Appendices I and II). Of the seven populations examined, these four had high numbers of alleles per locus, polymorphism and observed mean heterozygosity. Populations from Yata, Nishi-asahigaoka and Kakitagawa had relatively high percentages of soybean coverage, from 40% to 80% and

had low amounts of genetic variation.

Unfortunately, neither the distribution of the soybean plants (i.e., whether the plants were spatially scattered or clumped within each collection site) nor the types and distributions of co-habitating plant species were recorded in all instances. An additional problem is that the sizes of the collection areas were quite variable (Appendix I), confounding the estimate of wild soybean coverage. It is not appropriate, therefore, to speculate on the exact nature of the intra- and inter-specific plant associations or their effects on the population structure of the wild soybeans.

An additional consideration to the interpretation of the results in this study is that the ages of the wild soybean populations relative to one another are not known. The age of a population and the number of population founders will have a profound impact on the amounts of genetic variation present at any one time (Levin, 1975; Moran and Marshall, 1978; Warwick et al., 1987). If significant differences do exist among the wild soybean populations with respect to age, the interpretation of the observed patterns of genetic variation presented here may be flawed.

# CHAPTER III

## QUANTITATIVE TRAIT VARIATION WITHIN AND AMONG NATURAL POPULATIONS OF *GLYCINE SOJA*

## INTRODUCTION

Our current conceptual framework of the nature of quantitative variation in plant populations is based largely on the classic reciprocal transplant studies of Turesson (1922) and of Clausen et al. (1940, 1948). Three broad generalizations emerged from their research and research that followed (Hesslop-Harrison, 1964). First, widespread phenotypic variation exists within and among most plant populations. Second, phenotypic variation often is highly correlated with habitat variation. Finally, phenotypic variation not attributable to a plastic response often is the result of adaptation to local environmental conditions through genetic change. Turreson (1922) coined the word "ecotypes" to describe localized races genetically adapted to a particular environmental regime. Wright (1951) later suggested that such localized differentiation may be an important part of maintaining genetic variation and improving population fitness in the species as a whole.

In the past two decades, our understanding of plant population biology has advanced even further. We now know that genetically based morphological differences among

populations can be maintained even in the face of extensive gene flow (Antonovics et al., 1971; Jain and Bradshaw, 1966; Kiang, 1982) and can occur rapidly (Snaydon and Davies, 1972; Rai and Jain, 1982). Stochastic forces, such as genetic drift and founder events, also are recognized as processes which can contribute to differentiation among populations (Rai and Jain, 1982). The effective population size in plants (i.e. the number of inter-breeding members of a population) may be very small, as few as one to three individuals in an area of only a few square meters (Jain and Bradshaw, 1966; Snaydon and Davies, 1972; Linhart, 1974). Small effective population sizes ($N_e$) are especially common among predominantly self-pollinating plants such as Hordeum (barley) and Avena (slender oat) (Allard et al., 1968; Rai and Jain, 1982). Finally, we have a finer appreciation for phenotypic plasticity as a phenonomenon that has a genetic component and, therefore, is selectable by the environment (Schlichting, 1986). The precise mechanisms of phenotypic plasticity, however, have remained elusive to researchers.

## Experimental approaches to exploring phenotypic variation

Although morphological variation usually has the advantage of being easily discernable, the genetic basis of this variation often is complicated and not amenable to simple genetic interpretation. Expression of quantitative traits often is conditioned by a large number of gene loci

as well as additive, dominance and epistatic effects (Schwaegerle, et al., 1986). Epigenetic responses, such as phenotypic plasticity, also are common in plant systems (Schlichting, 1986). A standard method for separating genetic from environment-based effects in phenotypic variation is to grow plants in a common environment. Presumably, phenotypic differences between individuals under such conditions will be due largely to genetic differences. The randomization of plant choice and arrangement in the "common environment" gives reasonable assurance of achieving representative measurements of phenotypic characters (Greenlee and Rai, 1986).

The "common garden" approach has several shortcomings. First, it does not address directly the issue of phenotypic plasticity. To do so would require measuring characters under a variety of environmental conditions and/or transplant studies (Schlichting, 1986). Another drawback is that differences uncovered among individuals in a "common garden" design may have little relevance to plants in natural settings (Wu and Jain, 1978; Primack and Antonovics, 1981). Finally, the assumption of uniformity in a common environment, whether it is a field plot, greenhouse or growth chamber, often is flawed.

In this chapter, I describe the results of a two-year study on quantitative variation within and among seven natural populations of G. soja. The purposes of the study

were to 1) quantify quantitative variation of wild soybean plants grown in a common greenhouse environment, 2) identify suites of traits most useful in describing population differentiation using multivariate statistical analyses and 3) explore associations between phenotypic variation and environmental variation.

## MATERIALS AND METHODS

### A. Experimental methodology

Thirty-two traits (phenological, morphological and agronomic) were examined for wild soybean plants grown in the greenhouse during 1986 and 1987. Seeds were planted in late May, and the same greenhouse location was used in both years to minimize environmental variation. Field studies were not feasible given the long maturity of the wild soybean and the relatively short growing season in New Hampshire (Chiang, 1985).

1. Planting design (1986 and 1987). One seed per parent plant from each of 15 parent plants per population was grown using a completely randomized design in the greenhouse. Only one seed per parent plant was used due to the limited number of seeds originally collected from the field. If a seed did not germinate, it was eliminated from the analysis. Five of the 15 plants were sacrificed at 10 weeks to obtain root nodule and dry weight data. Two exceptions to this planting design were made for populations 6 (Asahigaoka-2) and 7 (Kakitagawa). Only 9 and 3 plants, respectively, were collected from these areas. For Asahigaoka-2, 14 plants were grown (5 were sacrificed at 10 weeks). For Kakitagawa, six plants were grown (3 were sacrificed at 10 weeks).

2. Growth conditions. Seeds were scarified, inoculated with a commercial strain of Rhizobium japonicum and planted at a depth of 1-2 cm in 8 inch (22 cm) diameter clay pots. A 1:1 mixture of sterilized field soil and Promix (sphagnum moss and vermiculite) was used as the growth medium. Pots were placed on the benches at 45 cm intervals. Bamboo stakes (120 cm long) were placed in the pots once the seedlings had germinated so that the plants could climb up the stakes and not become entangled. Plants were grown under natural light conditions. Insecticides were applied as needed to control two-spotted spider mites (Tetranychus urticae) and white flies (Trialeurodes vatorariorum).

3. Traits examined. The traits examined in the present study were selected based on previous studies of Glycine soja (Chiang, 1985). The following traits were examined for each plant:

Phenological traits

1. Number of days from sowing to germination (DTG)
2. Number of days from germination to first flower (DTF)
3. Number of days between first flower and first green pod (DFP)
4. Number of days between first green pod and first mature pod (DDP)
5. Number of days from the first dry pod to the last dry pod (LDP)
6. Number of days from germination to the last dry pod harvested (life span) (LSN)

Morphological traits

7. Number of branches at 4 weeks (B4W)
8. Number of leaves at 4 weeks (L4W)
9. Height of plant at 4 weeks (cm) (H4W)

10. Growth habit at 7 weeks (GH)
11. Number of leaves at 7 weeks (L7W)
12. Ovule number (ONO)
13. Banner petal width (cm) (BPW)
14. Flower length (cm) (FL)
15. Corolla tube length (cm) (CTL)
16. Length of central leaflet (cm) (LCL)
17. Width of central leaflet (cm) (WCL)
18. Ratio of length to width of central leaflet (LRC)
19. Length of lateral leaflet (cm) (LLL)
20. Width of lateral leaflet (cm) (WLL)
21. Ratio of length to width of lateral leaflet (LRL)
22. Root fluorescence (RF)

7-11. At four weeks post-germination, the number of branches (B4W) and leaves (L4W) were counted for all plants. Height (H4W), as measured from the cotyledon leaf scar to the apex, also was recored at this time. The growth habit (GH) of the soybean plants was classified as either "vine-like" or "compact" at 7 weeks (post-germination) based on the distance between the first and second growth nodes. Plants with internodal distances greater than 0.5 cm were considered vine-like and were assigned the value, 1. These plants had at least one or more branches that wrapped around the bamboo stake and grew upward rather than outward. Plants with internodal distances of 0.5 cm or less were considered compact and were assigned the value, 2. These plants had a 'bushy' appearance and had a tendency to extend branches outward instead of upward. By ten weeks, all plants had branches that were growing up the bamboo stake; by the end of the growing season, the plants were indistinguishable with respect to growth habit. Height was not measured at seven weeks because many of the soybean plants had branches

which were wrapped tightly around the bamboo stakes and could not, therefore, be measured accurately. The number of leaves was counted for each plant at seven weeks (L7W).

12. Ovule number was determined by dissecting a random sample of 10 flower buds per plant under a dissecting microscope and then counting the number of ovules present.

13-15. Ten (10) random flowers from each plant were measured one week after flowering began. The width of the banner petal (BPW), length of the corolla tube (CTL) and distance from the tip of the banner petal to the bottom of the corolla tube (FL) were measured (in centimeters) using a small plastic metric ruler under a dissecting microscope.

16-21. Three random trifoliate leaves were sampled from the top, middle and bottom third of the canopy as determined by the approximate height of the plant at time of flowering. The lengths and widths of the central leaflets were recorded (LCL and WCL, respectively). Because the lateral leaflets were virtually identical in size and shape, measurements on only one of the lateral leaflets were used in the analysis (LLL and WLL). The ratios of the lengths to widths of the central and lateral leaflets also were calculated (LRC and LRL, respectively) to reflect leaf shape.

22. Root fluorescence was determined by germinating 25 seeds per plant on moist paper towels in the dark and then examining the emerging radical using a UV (330 nm) light

source. Because the number of seeds originally collected from the wild soybean plants in the field was limited, progenies of the original wild soybean plants were used. Minsoy, a G. max cultivar known to be negative for root fluorescence, was used as a control (Delannay and Palmer, 1982).

Agronomic traits

23. Above ground dry weight at 10 weeks (gm) (AGW)
24. Below ground dry weight at 10 weeks (gm) (BGW)
25. Nodule dry weight at 10 weeks (gm) (NDW)
26. Number of root nodules at 10 weeks (NRN)
27. Total number of pods per plant (PPP)
28 a-d. Proportion of 1-, 2-, 3- and 4-seeded pods (P1SP, P2SP, P3SP, P4SP)
29. Seed weight per plant (gm) (SWP)
30. Seed weight per 100 seeds (gm) (SWH)
31. Total number of seeds per plant (SPP)
32. Average number of seeds per pod (ASP)

23-26. To obtain the data for above ground weight (AGW), below ground weight (BGW), number of root nodules (NRN) and nodule dry weight (NDW), 5 plants per population (3 plants for Kakitagawa) were sacrificed at 10 weeks post-germination. The above-ground portion of the plant was clipped off at the soil surface and placed, leaves and all, into a labelled brown paper bag for drying. Individual root systems were soaked for 2 hours in buckets of water to loosen the soil attached to the root fibers. The soil then was gently washed off of the root systems using a spray nozzle attached to a garden hose. Because the number of root nodules was too great to be counted accurately, an 8 cm diameter circle (masking tape spool) was centered over the

remaining portion of the stem and all nodules within the circle were counted. The nodules and the root systems were labelled and placed into brown bags for drying. All plant materials were dried for seven days at 49-54° C.

27-28. Mature dry pods were harvested daily from all plants until no more were produced by the plant. The total number of pods per plant (PPP), as well as the proportion of 1-, 2-, 3- and 4- seeded pods (P1SP, P2SP, P3SP, P4SP), was recorded.

29-32. The seeds were separated from the pods harvested from each plant and weighed (SWP). Seed weight per plant (SWP) and the weight of 100 random seeds (SWH) per plant were used to estimate the number of seeds per plant (SPP). The average number of seeds per pod (ASP) was calculated as the total number of pods harvested per plant divided by the estimated number of seeds per plant (PPP/SPP).

## B. Statistical analyses (1986 and 1987)

Morphological characters often are correlated, complicating statistical analysis and interpretation of phenotypic variation within and among populations (Lande and Arnold, 1983; Greenlee and Rai, 1986; Welsh et al., 1988); therefore, both univariate and multivariate approaches were used to analyze the phenotype data. A summary of the variables used in the different kinds of analyses is given in Appendix VII.

**Univariate analyses**

1. Analysis of variance (ANOVA). A one-way ANOVA for unbalanced data sets was performed on each of the quantitative characters to test for differences among population means (PROC GLM; SAS Institute, 1985b).

Assumptions for ANOVA include linearity, normality and homoscedasticity of error variances (Sokal and Rohlf, 1969). Linearity and normality were assessed by visual examination of histograms for each variable's distribution over all seven soybean populations. The plots were generated using the FREQUENCIES subroutine in SPSSX (SPSSX Inc., 1986). Homoscedasticity was assessed by Bartlett's test (Sokal and Rohlf, 1969) using the ONEWAY subroutine in SPSSX (SPSSX Inc., 1986). Data transformations were performed as necessary following the guidelines of Little and Hills (1978).

In instances where error variances could not be normalized or stabilized with standard transformations (e.g. square root, logarithmic, arcsin), the seriousness of the violations of assumptions was checked by comparing the results of the ANOVA to the results of a non-parametric Kruskal-Wallis test. Non-parametrics are concerned only with the distribution of variables, not with specific statistical parameters (such as the mean in ANOVA). The Kruskal-Wallis test compares all groups simultaneously and determines if

they are or are not samples drawn from the same population (Sokal and Rohlf, 1969). If the results for a specific variable were the same for the ANOVA and the Kruskal-Wallis test, it was assumed that any violations to the assumptions for the analysis of variance were minor and the untransformed variable was retained. If the results were discordant, the variable was eliminated from further analyses as a measure of conservatism.

2. Spearman rank correlations. Spearman rank correlation coefficients were used to explore relationships 1) between the population means of the quantitative traits and the environmental data for the wild soybean seed collection sites (Appendices I and II) and 2) between Mahalanobis' distance and the environmental variables. The coefficients were generated using PROC CORR in SAS (SAS Institute, 1985a).

The Spearman rank correlation ($r_s$) is a non-parametric test which determines if two variables are related and specifies the degree of relationship (Sokal and Rohlf, 1969). It was used in the present study due to the non-normal distributions of the environmental variables and the values of $D^2$.

Multivariate analyses

1. Canonical discriminant analysis (CDA). CDA was used

to identify the most parsimonious suite of phenotypic and life-history characters which had the greatest capacity to discriminate among the seven natural populations of the wild soybean. The details of this analysis are described in Chapter II.

The assumptions for a discriminant analysis are that the data follow a multivariate normal distribution and that the variance-covariance matrices are homogeneous. Variables which are highly correlated can lead to multicollinearity and singularity and, therefore, prevent matrix inversion, a necessary mathematical operation in multivariate analyses (Tabachnik and Fidel, 1983). Thus, variables which were linear combinations of other variables (i.e. P1SP, P2SP, P3SP and P4SP) and derived variables which were ratios of other variables (i.e., ASP, LRC and LRL) were excluded from discriminant analysis.

2. Mahalanobis' distance ($D^2$). Mahalanobis' distance, $D^2$, is a generalized measure of the distance between two groups in multivariate space based on the total covariance matrix of a particular data set (Tabachnik and Fidel, 1983). Population pairs with a distance of 0 are not different phenotypically. The phenotypic distance between two populations, a and b, is defined as

$$D_{ab}^2 = (n-g) \sum_{i=1}^{p} \sum_{j=1}^{p} w_{ij} * (\bar{X}_{ia} - \bar{X}_{ib})(\bar{X}_{ja} - \bar{X}_{jb}),$$

where g is the number of populations, p is the number of variables, n is the population size, wij* is an element of the inverse of the within-groups covariance matrix, $\overline{X}$ia and $\overline{X}_{i\,b}$ are the means for the ith variable in populations a and b, respectively, and $\overline{X}_{j\,a}$ and $\overline{X}_{j\,b}$ are the means for the jth variable in populations a and b, respectively (Norusis, 1985). $D^2$ was generated using PROC CANDISC in SAS (SAS Institute, 1985b).

# RESULTS

## A. Missing data

1. 1986. Two seeds, one from population 5 (Asahigaoka-1) and one from population 6 (Asahigaoka-2), did not germinate. Thus, the sizes of these populations were reduced from 15 to 14 and from 14 to 13, respectively.

2. 1987. One plant from population 4 (Yata) died shortly after germination, reducing the total number of plants examined for this population from 15 to 14.

## B. Data transformations

Skewed distributions and heteroscedasticity were detected for a number of variables in both the 1986 and 1987 data sets. Data transformations normalized and stabilized error variances in most instances. For a small number of variables, however, none of the standard data transformations corrected the violations of assumptions. For these variables, the results of the ANOVA and the non-parametric Kruskal-Wallis test were compared. If the results of the tests were the same, the variable was retained in its untransformed state. If the results were discordant, the variable was eliminated from further analyses. A summary of the data transformations used in this study is given in Appendix VII.

C. Within-population variation for quantitative traits

The means, standard errors, coefficients of variation and ranges of the quantitative characters examined for the seven natural populations of wild soybean in 1986 and 1987 are given in Tables 12 (pages 117-128) and 13 (pages 129-140).

Population 4 (Yata) was the most phenotypically distinct among all the populations of wild soybean examined. All individuals within population 4 had vine-like growth, whereas all other populations were composed of individuals of both compact and vine-like growth types. Flower and leaf sizes generally were larger for individuals from Yata than for individuals from other populations. In 1986, leaf shape in population 4 was clearly more ovate than in the other populations; in 1987, this distinction was blurred somewhat because leaf length was not significantly different among the populations. Leaf width, however, was greatest in population 4 in 1987. With respect to yield characters, soybean plants from Yata produced the least number of pods and seeds of any population, but had the highest seed weight per 100 seeds and the greatest proportion of 4-seeded pods.

Population 3 (Nishi-asahigaoka) and, to some extent, population 7 (Kakitagawa) also were distinct. These populations had individuals with both vine-like and compact growth habits, but had proportionally greater numbers of

vine-like individuals. Compared to all other populations, they had among the smallest flowers and leaves. Plants from populations 3 and 7 produced more seeds and pods than plants from population 4, but fewer than plants from the other four populations (1, 2, 5 and 6).

Quantitative trait means from populations 1, 2, 5 and 6 were all quite similar. Flower and leaf sizes were intermediate to those for individuals from populations 4 and populations 3 and 7. Plants from populations 1, 2, 5 and 6 produced much larger numbers of pods and seeds than plants from populations 3, 4 and 7. Average seed weight per 100 seeds for these four populations was noticeably less than for population 4, but similar to populations 3 and 7.

Year-to-year differences were detected in population means for most of the quantitative traits examined (Tables 12 and 13). The phenological traits, in particular, were increased (in units of days) in 1987 compared to 1986. Although the plants were grown in the same greenhouse location both years, fluctuations in the weather conditions probably affected the expression of the quantitative traits. In addition, the plants grown in 1986 and 1987 were started from seed which had been collected in 1982-1983 and then stored in the freezer (0 °C) at UNH. Thus, the seed grown in 1987 was a year older than that used in 1986 and some deterioration of seed quality may have occurred.

Table 12. Means ($\bar{X}$), standard errors (SE), coefficients of variation (CV) and ranges of phenotypic and life-history traits (untransformed data) for seven natural populations of the wild soybean. Data were collected from plants grown in the greenhouse in 1986.

| | Character[a] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1. Days to Germination (DTG) | | | | 2. Days to First Flower (DTF) | | | | 3. Days to First Pod (DFP) | | | |
| Population[b] | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 4.2 | 0.20 | 15.1 | 4-6 | 103.0 | 1.9 | 5.9 | 94-112 | 18.0 | 1.4 | 23.8 | 13-25 |
| 2 | 4.3 | 0.21 | 15.9 | 4-6 | 112.2 | 2.6 | 7.2 | 97-120 | 10.6 | 1.6 | 49.4 | 6-21 |
| 3 | 4.1 | 0.23 | 18.0 | 3-6 | 114.6 | 2.0 | 5.6 | 107-124 | 9.6 | 0.9 | 30.4 | 5-13 |
| 4 | 4.1 | 0.10 | 4.1 | 4-5 | 112.6 | 1.5 | 4.1 | 105-120 | 11.2 | 0.7 | 20.5 | 8-15 |
| 5 | 4.3 | 0.17 | 11.5 | 4-5 | 110.2 | 3.9 | 10.5 | 94-130 | 13.6 | 2.1 | 45.5 | 5-25 |
| 6 | 4.6 | 0.18 | 16.1 | 4-6 | 113.5 | 2.5 | 6.3 | 101-121 | 13.2 | 1.6 | 33.9 | 8-19 |
| 7 | 4.3 | 0.33 | 13.3 | 4-5 | 110.3 | 7.8 | 12.3 | 95-121 | 14.0 | 5.6 | 68.9 | 7-25 |
| Overall mean | 4.3 | 0.08 | 14.2 | 3-6 | 110.9 | 1.08 | 7.6 | 94-130 | 12.7 | 0.7 | 41.2 | 5-25 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 4. Days to First Dry Pod (DDP) $\bar{X}$ | SE | CV | Range | 5. Days to Last Dry Pod (LDP) $\bar{X}$ | SE | CV | Range | 6. Life Span (LSN) $\bar{X}$ | SE | CV | Range |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 26.5 | 1.2 | 14.4 | 23-34 | 27.8 | 1.9 | 21.1 | 21-42 | 183.8 | 3.0 | 5.4 | 173-206 |
| 2 | 28.0 | 1.2 | 13.3 | 22-32 | 27.0 | 0.95 | 11.2 | 24-34 | 187.7 | 2.0 | 3.4 | 178-196 |
| 3 | 26.4 | 1.1 | 13.1 | 22-34 | 26.9 | 0.74 | 8.7 | 23-30 | 187.2 | 2.4 | 4.1 | 179-199 |
| 4 | 27.6 | 1.3 | 14.4 | 22-36 | 25.0 | 0.71 | 9.0 | 21.28 | 186.3 | 1.9 | 3.3 | 180-195 |
| 5 | 24.6 | 1.3 | 16.3 | 18-31 | 26.7 | 0.76 | 8.6 | 22-29 | 188.2 | 3.3 | 5.3 | 174-201 |
| 6 | 28.2 | 2.4 | 23.6 | 19-39 | 27.4 | 2.3 | 23.6 | 21-42 | 193.6 | 3.7 | 5.4 | 180-208 |
| 7 | 29.7 | 1.2 | 7.0 | 28-32 | 26.3 | 0.88 | 5.8 | 25-28 | 190.0 | 3.5 | 3.2 | 184-196 |
| Overall mean | 27.0 | 0.55 | 15.7 | 18-39 | 26.7 | 0.49 | 14.2 | 21-42 | 187.8 | 1.1 | 4.5 | 173-208 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 7. Number of Branches at Four Weeks (B4W) | | | | 8. Number of Leaves at Four Weeks (L4W) | | | | 9. Height at Four Weeks (cm) (H4W) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 3.0 | 0.36 | 38.5 | 1-4 | 8.9 | 0.90 | 31.9 | 5-14 | 2.04 | 0.23 | 34.9 | 1- 3 |
| 2 | 2.2 | 0.20 | 28.7 | 2-4 | 7.9 | 0.10 | 4.0 | 7- 8 | 1.09 | 0.19 | 54.1 | 0.6- 2.4 |
| 3 | 2.1 | 0.31 | 47.4 | 1-4 | 7.9 | 0.69 | 27.6 | 4-11 | 9.66 | 3.3 | 106.9 | 0.1- 2.6 |
| 4 | 2.9 | 0.31 | 34.3 | 2-4 | 8.8 | 0.42 | 14.9 | 8-11 | 4.64 | 0.58 | 39.2 | 1.3- 6.7 |
| 5 | 2.8 | 0.32 | 34.9 | 2-4 | 10.7 | 1.1 | 29.6 | 8-14 | 2.34 | 0.60 | 76.4 | 1.2- 6.8 |
| 6 | 2.3 | 0.31 | 39.4 | 1-4 | 8.8 | 1.4 | 43.9 | 2-14 | 2.26 | 0.96 | 119.5 | 0.9- 8.9 |
| 7 | 2.0 | 0 | 0 | 2 | 9.7 | 2.2 | 39.2 | 7-14 | 2.66 | 1.7 | 112.3 | 0.6- 6.1 |
| Overall mean | 2.5 | 0.12 | 38.4 | 1-4 | 8.8 | 0.33 | 29.3 | 7-14 | 3.69 | 0.67 | 140.5 | 0.1-26.0 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 10. Growth Habit (GH) | | | | 11. Number of Leaves at Seven Weeks (L7W) | | | | 12. Ovule number (ONO) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 1.5 | 0.17 | 35.1 | 1-2 | 70.7 | 4.9 | 22.1 | 41-95 | 2.96 | 0.03 | 3.4 | 2.8-3.1 |
| 2 | 1.9 | 0.10 | 16.6 | 1-2 | 70.7 | 2.8 | 12.4 | 59-89 | 3.03 | 0.04 | 3.8 | 2.9-3.3 |
| 3 | 1.2 | 0.13 | 35.1 | 1-2 | 63.1 | 6.0 | 30.2 | 32-92 | 2.95 | 0.02 | 2.4 | 2.8-3.0 |
| 4 | 1.0 | 0 | 0 | 1 | 65.2 | 8.5 | 41.2 | 1-92 | 3.08 | 0.03 | 3.4 | 3.0-3.3 |
| 5 | 1.6 | 0.18 | 33.9 | 1-2 | 75.6 | 3.9 | 15.3 | 56-89 | 3.06 | 0.03 | 2.9 | 2.9-3.2 |
| 6 | 1.6 | 0.18 | 31.8 | 1-2 | 60.5 | 8.2 | 38.4 | 8-80 | 3.03 | 0.04 | 3.6 | 2.9-3.2 |
| 7 | 1.7 | 0.33 | 34.6 | 1-2 | 44.7 | 13.5 | 52.5 | 19-65 | 3.00 | 0.06 | 3.3 | 2.9-3.1 |
| Overall mean | 1.5 | 0.06 | 34.3 | 1-2 | 66.6 | 2.5 | 28.7 | 1-95 | 3.02 | 0.01 | 3.5 | 2.8-3.3 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 13. Banner Petal Width (cm) (BPW) | | | | 14. Flower Length (cm) (FL) | | | | 15. Corolla Tube Length (cm) (CTL) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 5.4 | 0.05 | 2.6 | 5.1-5.6 | 6.2 | 0.07 | 3.3 | 5.7-6.4 | 2.3 | 0.03 | 4.2 | 2.1-2.4 |
| 2 | 5.5 | 0.06 | 3.5 | 5.1-5.7 | 6.2 | 0.02 | 1.2 | 6.2-6.4 | 2.3 | 0.02 | 2.4 | 2.2-2.4 |
| 3 | 4.4 | 0.06 | 4.2 | 4.2-4.7 | 6.2 | 0.04 | 2.0 | 6.0-6.4 | 2.3 | 0.02 | 2.1 | 2.2-2.4 |
| 4 | 5.8 | 0.04 | 2.3 | 5.5-6.0 | 7.0 | 0.05 | 2.4 | 6.8-7.3 | 2.4 | 0.06 | 7.3 | 2.3-2.9 |
| 5 | 5.6 | 0.03 | 1.8 | 5.5-5.8 | 6.5 | 0.06 | 2.9 | 6.3-6.8 | 2.2 | 0.03 | 3.6 | 2.1-2.4 |
| 6 | 5.5 | 0.06 | 3.3 | 5.3-5.8 | 6.3 | 0.08 | 3.4 | 5.9-6.6 | 2.3 | 0.03 | 3.9 | 2.3-2.5 |
| 7 | 4.9 | 0.04 | 1.6 | 4.8-4.9 | 6.5 | 0.10 | 2.7 | 6.3-6.6 | 2.3 | 0.04 | 3.3 | 2.2-2.4 |
| Overall mean | 5.3 | 0.06 | 8.9 | 4.2-6.0 | 6.4 | 0.05 | 5.6 | 5.7-7.3 | 2.3 | 0.01 | 4.9 | 2.1-2.9 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population | Character[a] 16. Length of the Central Leaflet (cm) (LCL) | | | | 17. Width of the Central Leaflet (cm) (WCL) | | | | 18. Ratio of Length to Width of Central Leaflet (LRC) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$[b] | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 7.1 | 0.15 | 6.6 | 6.4-7.9 | 3.1 | 0.09 | 9.9 | 2 7-3.6 | 2.3 | 0.06 | 8.3 | 2.0-2.6 |
| 2 | 6.0 | 0.18 | 9.3 | 5.1-6.9 | 2.8 | 0.08 | 9.1 | 2.4-3.2 | 2.1 | 0.06 | 8.5 | 1.9-2.4 |
| 3 | 6.2 | 0.11 | 5.4 | 5.7-6.9 | 3.3 | 0.06 | 6.1 | 3.0-3.6 | 1.9 | 0.04 | 7.2 | 1.7-2.1 |
| 4 | 6.7 | 0.16 | 7.8 | 6.0-7.6 | 4.1 | 0.12 | 9.5 | 3.6-4.8 | 1.6 | 0.03 | 4.9 | 1.5-1.8 |
| 5 | 6.4 | 0.27 | 11.9 | 5.3-7.5 | 3.0 | 0.23 | 21.4 | 1.8-3.7 | 2.2 | 0.13 | 16.8 | 1.9-3.0 |
| 6 | 6.1 | 0.34 | 15.7 | 4.6-7.2 | 3.1 | 0.26 | 23.6 | 2.2-4.4 | 2.0 | 0.14 | 19.7 | 1.3-2.7 |
| 7 | 5.8 | 0.05 | 1.5 | 5.7-5.9 | 3.0 | 0.09 | 5.1 | 2.8-3.1 | 1.9 | 0.06 | 5.4 | 1.8-2.1 |
| Overall mean | 6.4 | 0.09 | 10.9 | 4.6-7.9 | 3.2 | 0.08 | 18.1 | 1.8-4.8 | 2.0 | 0.04 | 15.8 | 1.3-3.0 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 19. Length of the Lateral Leaflet (cm) (LLL) | | | | 20. Width of the Lateral Leaflet (cm) (WLL) | | | | 21. Ratio of Length to Width of the Lateral Leaflet (LRL) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 5.6 | 0.13 | 7.3 | 4.9-6.2 | 2.9 | 0.13 | 13.9 | 2.1-3.5 | 1.9 | 0.07 | 11.7 | 1.8-2.6 |
| 2 | 4.8 | 0.13 | 8.8 | 4.0-5.4 | 2.9 | 0.09 | 9.9 | 2.4-3.3 | 1.7 | 0.04 | 7.0 | 1.5-1.9 |
| 3 | 4.9 | 0.09 | 5.5 | 4.6-5.5 | 3.0 | 0.07 | 7.5 | 2.7-3.4 | 1.6 | 0.03 | 5.9 | 1.5-1.8 |
| 4 | 5.7 | 0.15 | 8.4 | 5.0-6.5 | 3.8 | 0.13 | 10.9 | 3.2-4.5 | 1.5 | 0.03 | 5.4 | 1.4-1.7 |
| 5 | 5.1 | 0.24 | 13.2 | 4.2-5.9 | 3.1 | 0.22 | 19.9 | 1.9-3.7 | 1.7 | 0.07 | 11.6 | 1.6-2.2 |
| 6 | 5.0 | 0.30 | 16.9 | 3.7-5.7 | 3.0 | 0.22 | 20.2 | 2.2-3.9 | 1.7 | 0.06 | 10.9 | 1.5-2.0 |
| 7 | 4.7 | 0.15 | 5.6 | 4.5-5.0 | 3.5 | 0.62 | 30.8 | 2.8-4.7 | 1.4 | 0.18 | 21.7 | 1.1-1.6 |
| Overall mean | 5.2 | 0.08 | 11.8 | 3.7-6.5 | 3.5 | 0.62 | 30.8 | 2.8-4.7 | 1.7 | 0.03 | 12.8 | 1.1-2.6 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 23. Above-Ground Weight (gm) (AGW) | | | | 24. Below-Ground Weight (gm) (BGW) | | | | 25. Nodule Dry Weight (gm) (NDW) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 16.2 | 2.3 | 32.2 | 12.1-24.5 | 12.8 | 3.3 | 57.7 | 4.5-21.3 | 0.40 | 0.05 | 30.8 | 0.26-0.55 |
| 2 | 15.4 | 2.6 | 37.0 | 10.6-25.2 | 10.8 | 1.7 | 35.9 | 5.1-15.4 | 0.35 | 0.05 | 31.6 | 0.22-0.47 |
| 3 | 20.5 | 2.3 | 25.4 | 11.3-23.5 | 10.1 | 2.1 | 46.4 | 4.6-17.1 | 0.55 | 0.14 | 55.4 | 0.03-0.77 |
| 4 | 22.5 | 1.2 | 11.6 | 10.0-24.3 | 11.1 | 0.89 | 17.8 | 9.4-14.1 | 0.62 | 0.07 | 24.0 | 0.48-0.87 |
| 5 | 16.9 | 1.9 | 25.9 | 12.0-22.7 | 13.3 | 3.1 | 52.3 | 7.7-24.9 | 0.42 | 0.05 | 27.2 | 0.30-0.59 |
| 6 | 19.4 | 3.4 | 38.9 | 14.7-32.5 | 10.2 | 2.0 | 44.5 | 6.8-17.9 | 0.41 | 0.05 | 29.3 | 0.30-0.60 |
| 7 | 17.1 | 3.3 | 33.7 | 12.7-23.6 | 8.0 | 2.0 | 43.8 | 4.8-11.8 | 0.42 | 0.16 | 66.2 | 0.24-0.74 |
| Overall mean | 18.4 | 0.94 | 29.5 | 10.6-32.5 | 11.1 | 0.85 | 43.9 | 4.5-24.9 | 0.45 | 0.03 | 41.0 | 0.03-0.87 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 26. Number of Root Nodules (NRN) | | | | 27. Number of Pods per Plant (PPP) | | | | 28a. Proportion of One-Seed Pods (P1SP) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 106.0 | 16.6 | 35.1 | 55-157 | 1182.9 | 56.2 | 15.0 | 947-1442 | 0.10 | 0.01 | 43.6 | 0.04-0.18 |
| 2 | 117.0 | 17.5 | 33.4 | 74-170 | 1048.0 | 70.1 | 21.1 | 856-1443 | 0.05 | 0.01 | 68.0 | 0.02-0.15 |
| 3 | 160.8 | 40.7 | 56.7 | 8-232 | 958.7 | 52.1 | 17.2 | 699-1247 | 0.13 | 0.01 | 34.7 | 0.06-0.24 |
| 4 | 172.8 | 33.4 | 43.2 | 81-271 | 757.5 | 22.7 | 9.5 | 644- 864 | 0.08 | 0.01 | 33.3 | 0.04-0.11 |
| 5 | 96.4 | 4.6 | 10.8 | 78-103 | 1046.3 | 46.6 | 13.4 | 896-1315 | 0.10 | 0.02 | 53.4 | 0.03-0.20 |
| 6 | 119.6 | 26.5 | 49.6 | 70-221 | 1016.5 | 67.3 | 18.7 | 731-1288 | 0.06 | 0.01 | 36.9 | 0.03-0.10 |
| 7 | 149.7 | 60.8 | 70.3 | 70-269 | 980.0 | 75.5 | 13.3 | 845-1106 | 0.07 | 0.03 | 81.2 | 0.02-0.13 |
| Overall mean | 130.7 | 11.0 | 48.6 | 8-271 | 999.3 | 26.3 | 20.4 | 644-1443 | 0.09 | 0.01 | 53.2 | 0.02-0.24 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population | Character[a] 28b. Proportion of Two-Seed Pods (P2SP) | | | | 28c. Proportion of Three-Seed Pods (P3SP) | | | | 28d. Proportion of Four-Seed Pods (P4SP) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$[b] | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 0.40 | 0.02 | 16.5 | 0.28-0.50 | 0.49 | 0.03 | 20.7 | 0.34-0.66 | 0.01 | 0.001 | 47.3 | 0 -0.02 |
| 2 | 0.28 | 0.03 | 35.1 | 0.20-0.47 | 0.66 | 0.03 | 14.9 | 0.40-0.76 | 0.04 | 0.01 | 75.5 | 0.002-0.09 |
| 3 | 0.38 | 0.01 | 11.4 | 0.29-0.45 | 0.48 | 0.02 | 14.2 | 0.39-0.60 | 0.01 | 0.002 | 52.8 | 0.003-0.02 |
| 4 | 0.30 | 0.01 | 15.5 | 0.21-0.36 | 0.56 | 0.02 | 9.3 | 0.50-0.64 | 0.07 | 0.01 | 30.7 | 0.04 -0.10 |
| 5 | 0.28 | 0.02 | 20.2 | 0.19-0.39 | 0.58 | 0.03 | 17.4 | 0.45-0.73 | 0.04 | 0.01 | 55.2 | 0.004-0.08 |
| 6 | 0.27 | 0.02 | 25.6 | 0.20-0.42 | 0.64 | 0.03 | 13.1 | 0.48-0.72 | 0.02 | 0.004 | 61.2 | 0.01 -0.04 |
| 7 | 0.18 | 0.02 | 20.8 | 0.14-0.22 | 0.70 | 0.04 | 10.9 | 0.62-0.77 | 0.04 | 0.01 | 46.4 | 0.03 -0.07 |
| Overall mean | 0.31 | 0.01 | 27.1 | 0.14-0.50 | 0.57 | 0.01 | 19.1 | 0.34-0.77 | 0.03 | 0.003 | 83.2 | 0 -0.10 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population | Character[a] 29. Seed Weight per Plant (gm) (SWP) | | | | 30. Seed Weight per One-Hundred Seeds (gm) (SWH) | | | | 31. Number of Seed per Plant (SPP) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$[b] | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 48.1 | 2.7 | 17.9 | 37.8-62.4 | 1.83 | 0.06 | 10.1 | 1.36-1.99 | 2630.6 | 124.5 | 14.9 | 2136-3320 |
| 2 | 48.1 | 2.9 | 18.9 | 37.6-63.4 | 1.79 | 0.04 | 6.3 | 1.64-2.03 | 2692.3 | 188.6 | 22.5 | 2224-3751 |
| 3 | 40.9 | 2.2 | 16.8 | 31.7-55.6 | 1.83 | 0.05 | 8.9 | 1.64-2.26 | 2249.2 | 133.0 | 18.7 | 1731-3161 |
| 4 | 44.3 | 1.5 | 10.7 | 36.6-52.2 | 2.43 | 0.06 | 8.4 | 2.07-2.67 | 1812.1 | 56.8 | 9.9 | 1551-2071 |
| 5 | 50.9 | 2.2 | 12.9 | 35.7-58.0 | 1.91 | 0.10 | 16.4 | 1.56-2.42 | 2722.0 | 172.7 | 19.0 | 2029-3320 |
| 6 | 45.4 | 2.3 | 14.5 | 34.7-53.3 | 1.92 | 0.09 | 12.9 | 1.67-2.47 | 2407.3 | 181.4 | 21.3 | 1790-3193 |
| 7 | 46.2 | 2.6 | 9.8 | 41.1-49.8 | 1.75 | 0.02 | 2.2 | 1.71-1.78 | 2637.3 | 177.1 | 11.6 | 2308-2915 |
| Overall mean | 46.2 | 0.97 | 16.3 | 31.7-63.4 | 1.94 | 0.04 | 15.3 | 1.36-2.67 | 2425.2 | 68.9 | 22.0 | 1551-3751 |

Table 12. 1986 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 32. Average number of Seeds per Pod (ASP) $\bar{X}$ | SE | CV | Range |
|---|---|---|---|---|
| 1 | 2.22 | 0.04 | 6.2 | 1.94-2.45 |
| 2 | 2.57 | 0.06 | 7.9 | 2.06-2.83 |
| 3 | 2.35 | 0.03 | 4.5 | 2.15-2.53 |
| 4 | 2.39 | 0.05 | 6.9 | 2.14-2.68 |
| 5 | 2.59 | 0.10 | 12.1 | 2.12-3.07 |
| 6 | 2.37 | 0.10 | 12.3 | 1.82-2.70 |
| 7 | 2.69 | 0.03 | 1.9 | 2.64-2.73 |
| Overall mean | 2.43 | 0.03 | 9.9 | 1.82-3.07 |

[a] For explanation of character notation, see Materials and Methods.

[b] Populations: 1= Hatsunedai; 2= Kanodanchi; 3= Nishi-asahigaoka; 4= Yata; 5= Asahigaoka-1; 6= Asahigaoka-2; 7= Kakitagawa.

Table 13. Means ($\bar{X}$), standard errors (SE), coefficients of variation (CV) and ranges of phenotypic and life-history traits (untransformed data) for seven natural populations of the wild soybean. Data were collected from plants grown in the greenhouse in 1987.

| | Character[a] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1. Days to Germination (DTG) | | | | 2. Days to First Flower (DTF) | | | | 3. Days to First Pod (DFP) | | | |
| Population[b] | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 6.8 | 0.79 | 11.6 | 5-8 | 110.7 | 1.5 | 4.3 | 108-120 | 15.1 | 0.46 | 9.6 | 13-17 |
| 2 | 6.9 | 0.23 | 10.7 | 6-8 | 123.4 | 1.6 | 4.0 | 111-127 | 11.3 | 0.56 | 15.6 | 10-15 |
| 3 | 6.8 | 0.13 | 6.2 | 6-7 | 119.7 | 0.9 | 2.5 | 113-124 | 13.6 | 0.92 | 21.4 | 10-20 |
| 4 | 6.8 | 0.15 | 6.5 | 6-7 | 121.0 | 1.1 | 2.8 | 115-127 | 13.6 | 0.91 | 20.2 | 9-18 |
| 5 | 6.9 | 0.23 | 10.7 | 6-8 | 119.9 | 2.1 | 5.5 | 110-127 | 13.0 | 0.92 | 22.4 | 9-17 |
| 6 | 7.1 | 0.26 | 11.0 | 6-8 | 121.4 | 1.8 | 4.6 | 111-127 | 11.6 | 0.73 | 18.7 | 10-17 |
| 7 | 6.7 | 0.33 | 8.7 | 6-7 | 123.0 | 2.0 | 2.8 | 121-127 | 10.3 | 2.30 | 39.1 | 6-14 |
| Overall mean | 6.9 | 0.08 | 9.4 | 5-8 | 119.5 | 0.79 | 5.2 | 108-127 | 12.9 | 0.35 | 21.1 | 6-20 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 4. Days to First Dry Pod (DDP) $\bar{X}$ | SE | CV | Range | 5. Days to Last Dry Pod (LDP) $\bar{X}$ | SE | CV | Range | 6. Life Span (LSN) $\bar{X}$ | SE | CV | Range |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 30.3 | 0.90 | 9.3 | 25-35 | 22.4 | 1.6 | 22.0 | 15-30 | 178.6 | 0.58 | 1.0 | 176-181 |
| 2 | 26.5 | 1.40 | 16.7 | 20-34 | 19.2 | 1.4 | 23.0 | 15-28 | 180.1 | 0.67 | 1.2 | 177-184 |
| 3 | 27.7 | 0.73 | 8.3 | 25-32 | 17.9 | 0.78 | 13.8 | 15-21 | 178.9 | 0.74 | 1.3 | 176-184 |
| 4 | 27.0 | 0.62 | 6.9 | 24-29 | 17.7 | 0.41 | 6.9 | 15-19 | 179.5 | 0.41 | 0.68 | 178-182 |
| 5 | 26.5 | 1.01 | 12.1 | 21-30 | 19.7 | 1.1 | 17.8 | 16-25 | 179.3 | 0.84 | 1.5 | 175-183 |
| 6 | 27.0 | 0.50 | 25.0 | 25-30 | 18.2 | 0.70 | 17.0 | 17-23 | 180.0 | 0.60 | 2.0 | 178-183 |
| 7 | 27.3 | 0.67 | 4.2 | 26-28 | 17.7 | 0.33 | 3.3 | 17-18 | 179.0 | 0.58 | 0.56 | 178-180 |
| Overall mean | 27.5 | 0.38 | 10.9 | 20-35 | 19.2 | 0.46 | 18.9 | 15-30 | 179.4 | 0.26 | 1.1 | 175-184 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 7. Number of Branches at Four Weeks (B4W) | | | | 8. Number of Leaves at Four Weeks (L4W) | | | | 9. Height at Four Weeks (cm) (H4W) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 3.6 | 0.27 | 23.4 | 2-4 | 11.3 | 0.70 | 19.6 | 8-14 | 2.7 | 0.42 | 48.1 | 1.3- 5.2 |
| 2 | 3.7 | 0.37 | 31.3 | 2-6 | 13.1 | 0.90 | 21.7 | 8-17 | 1.3 | 0.23 | 55.0 | 0.3- 3.0 |
| 3 | 2.8 | 0.44 | 49.9 | 2-6 | 10.6 | 1.3 | 38.6 | 8-18 | 5.3 | 2.1 | 122.7 | 1.5-21.2 |
| 4 | 3.8 | 0.22 | 17.6 | 2-4 | 10.7 | 0.33 | 9.4 | 8-11 | 4.8 | 0.46 | 28.7 | 2.4- 7.1 |
| 5 | 3.6 | 0.40 | 35.1 | 2-6 | 12.1 | 1.0 | 26.2 | 8-17 | 10.9 | 5.5 | 160.6 | 1.1-50.9 |
| 6 | 3.3 | 0.33 | 30.0 | 2-4 | 12.0 | 1.0 | 25.0 | 8-14 | 6.6 | 3.5 | 160.1 | 1.1-27.0 |
| 7 | 4.7 | 0.67 | 24.7 | 4-6 | 14.0 | 1.7 | 21.4 | 11-17 | 6.4 | 5.4 | 146.0 | 0.8-17.1 |
| Overall mean | 3.5 | 0.14 | 31.8 | 2-6 | 11.7 | 0.38 | 25.0 | 8-18 | 5.3 | 1.1 | 168.1 | 0.3-50.9 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 10. Growth Habit (GH) $\bar{X}$ | SE | CV | Range | 11. Number of Leaves at Seven Weeks (L7W) $\bar{X}$ | SE | CV | Range | 12. Ovule Number (ONO) $\bar{X}$ | SE | CV | Range |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.3 | 0.15 | 37.2 | 1-2 | 85.8 | 4.0 | 14.8 | 72-108 | 2.96 | 0.02 | 1.7 | 2.9-3.0 |
| 2 | 1.9 | 0.10 | 16.6 | 1-2 | 84.0 | 5.6 | 21.0 | 60-114 | 3.01 | 0.02 | 2.4 | 2.9-3.1 |
| 3 | 1.1 | 0.10 | 28.7 | 1-2 | 70.2 | 9.9 | 44.6 | 39-129 | 2.96 | 0.03 | 2.8 | 2.8-3.1 |
| 4 | 1.0 | 0 | 0 | 1 | 72.6 | 5.3 | 21.8 | 35- 90 | 3.02 | 0.04 | 4.2 | 2.9-3.3 |
| 5 | 1.5 | 0.17 | 35.1 | 1-2 | 81.6 | 6.0 | 23.2 | 54-117 | 3.03 | 0.03 | 2.9 | 2.9-3.2 |
| 6 | 1.7 | 0.17 | 30.0 | 1-2 | 85.7 | 5.8 | 20.4 | 60-120 | 2.98 | 0.03 | 2.8 | 2.9-3.1 |
| 7 | 1.7 | 0.17 | 34.6 | 1-2 | 86.0 | 8.7 | 17.6 | 72-102 | 3.00 | 0.06 | 3.3 | 2.9-3.1 |
| Overall mean | 1.4 | 0.06 | 34.9 | 1-2 | 80.3 | 2.5 | 24.7 | 35-129 | 2.99 | 0.01 | 2.9 | 2.8-3.3 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 13. Banner Petal Width (cm) (BPW) | | | | 14. Flower Length (cm) (FL) | | | | 15. Corolla Tube Length (cm) (CTL) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 5.0 | 0.10 | 6.3 | 4.6-5.7 | 6.1 | 0.06 | 3.1 | 5.6-6.3 | 2.0 | 0.01 | 1.9 | 1.9-2.1 |
| 2 | 5.0 | 0.04 | 2.3 | 4.8-5.2 | 6.1 | 0.04 | 2.3 | 5.7-6.2 | 2.0 | 0.01 | 0.9 | 2.0-2.1 |
| 3 | 4.2 | 0.09 | 6.7 | 3.6-4.6 | 5.6 | 0.10 | 5.5 | 5.1-6.1 | 2.0 | 0.01 | 2.4 | 1.9-2.1 |
| 4 | 5.1 | 0.12 | 6.4 | 4.7-5.8 | 6.0 | 0.14 | 6.7 | 5.5-6.8 | 2.0 | 0.02 | 2.2 | 1.9-2.1 |
| 5 | 5.0 | 0.12 | 7.5 | 4.7-5.9 | 6.1 | 0.10 | 5.4 | 5.7-6.7 | 2.0 | 0.01 | 1.4 | 2.0-2.1 |
| 6 | 5.2 | 0.18 | 10.5 | 4.8-6.4 | 6.2 | 0.17 | 8.2 | 5.6-7.1 | 2.0 | 0.02 | 2.6 | 1.9-2.1 |
| 7 | 4.6 | 0.08 | 3.0 | 4.5-4.7 | 5.8 | 0.11 | 3.3 | 5.6-6.0 | 2.0 | 0.01 | 0.8 | 2.0-2.1 |
| Overall mean | 4.9 | 0.58 | 9.3 | 3.6-6.4 | 6.0 | 0.05 | 6.0 | 5.1-7.1 | 2.0 | 0.01 | 1.9 | 1.9-2.1 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 16. Length of the Central Leaflet (cm) (LCL) | | | | 17. Width of the Central Leaflet (cm) (WCL) | | | | 18. Ratio of Length to Width of Central Leaflet (LRC) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 6.7 | 0.14 | 6.4 | 6.1-7.4 | 2.4 | 0.10 | 12.4 | 2.0-3.1 | 2.8 | 0.12 | 13.7 | 2.3-3.6 |
| 2 | 6.3 | 0.12 | 5.8 | 5.9-7.1 | 3.0 | 0.10 | 10.5 | 2.4-3.4 | 2.2 | 0.08 | 11.9 | 1.9-2.7 |
| 3 | 6.6 | 0.14 | 6.6 | 5.9-7.3 | 3.4 | 0.08 | 7.5 | 3.1-3.7 | 1.9 | 0.03 | 5.0 | 1.8-2.1 |
| 4 | 6.0 | 0.13 | 6.6 | 5.3-6.6 | 3.6 | 0.10 | 8.6 | 3.1-4.1 | 1.7 | 0.03 | 5.4 | 1.6-1.8 |
| 5 | 6.4 | 0.16 | 8.1 | 5.6-7.0 | 3.0 | 0.14 | 14.3 | 2.3-3.8 | 2.2 | 0.11 | 16.2 | 1.8-3.0 |
| 6 | 6.3 | 0.22 | 10.3 | 5.6-7.5 | 3.0 | 0.15 | 14.8 | 2.4-3.9 | 2.1 | 0.09 | 13.3 | 1.9-2.7 |
| 7 | 6.2 | 0.17 | 4.6 | 6.0-6.6 | 3.1 | 0.03 | 1.6 | 3.1-3.2 | 2.0 | 0.04 | 3.6 | 1.9-2.1 |
| Overall mean | 6.4 | 0.06 | 7.8 | 5.3-7.5 | 3.1 | 0.06 | 15.8 | 2.0-4.1 | 2.1 | 0.05 | 19.3 | 1.6-3.6 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 19. Length of the Lateral Leaflet (cm) (LLL) | | | | 20. Width of the Lateral Leaflet (cm) (WLL) | | | | 21. Ratio of Length to Width of the Lateral Leaflet (LRL) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 5.3 | 0.12 | 7.4 | 4.5-6.1 | 2.5 | 0.11 | 13.4 | 2.2-3.2 | 2.1 | 0.07 | 10.7 | 1.8-2.5 |
| 2 | 5.2 | 0.09 | 5.3 | 4.8-5.5 | 3.2 | 0.12 | 11.5 | 2.6-3.6 | 1.6 | 0.06 | 12.2 | 1.5-2.1 |
| 3 | 5.2 | 0.13 | 7.7 | 4.8-5.8 | 3.2 | 0.11 | 10.5 | 2.7-3.8 | 1.6 | 0.03 | 6.0 | 1.5-1.8 |
| 4 | 5.2 | 0.12 | 7.2 | 4.5-5.7 | 3.3 | 0.10 | 9.4 | 2.8-3.7 | 1.6 | 0.02 | 4.3 | 1.5-1.6 |
| 5 | 5.3 | 0.13 | 7.7 | 4.8-6.0 | 3.1 | 0.14 | 14.5 | 2.4-3.9 | 1.7 | 0.08 | 13.9 | 1.5-2.2 |
| 6 | 5.2 | 0.17 | 10.1 | 4.4-6.2 | 3.2 | 0.15 | 14.6 | 2.4-4.0 | 1.7 | 0.06 | 10.4 | 1.5-2.0 |
| 7 | 5.1 | 0.08 | 2.7 | 4.9-5.2 | 3.1 | 0.08 | 4.4 | 3.0-3.2 | 1.6 | 0.04 | 4.1 | 1.5-1.7 |
| Overall mean | 5.2 | 0.05 | 7.2 | 4.4-6.2 | 3.1 | 0.06 | 14.2 | 2.2-4.0 | 1.7 | 0.03 | 14.2 | 1.5-2.5 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 23. Above-Ground Weight (gm) (AGW) | | | | 24. Below-Ground Weight (gm) (BGW) | | | | 25. Nodule Dry Weight (gm) (NDW) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 17.2 | 2.8 | 36.7 | 10.7-26.7 | 6.2 | 1.4 | 51.6 | 2.7-10.8 | 0.35 | 0.09 | 58.2 | 0.14-0.57 |
| 2 | 14.9 | 1.1 | 16.7 | 11.0-17.3 | 7.1 | 0.84 | 26.7 | 4.1- 9.0 | 0.25 | 0.02 | 20.2 | 0.21-0.33 |
| 3 | 13.0 | 2.2 | 34.2 | 6.4-15.9 | 4.2 | 0.66 | 31.7 | 2.4- 5.6 | 0.19 | 0.08 | 80.1 | 0.05-0.37 |
| 4 | 21.7 | 1.8 | 17.1 | 18.2-25.9 | 6.5 | 0.42 | 12.8 | 5.8- 7.4 | 0.46 | 0.06 | 26.9 | 0.31-0.58 |
| 5 | 14.6 | 1.2 | 17.8 | 12.0-18.7 | 5.3 | 0.32 | 13.7 | 4.5- 6.2 | 0.21 | 0.02 | 19.9 | 0.17-0.28 |
| 6 | 15.6 | 1.4 | 20.3 | 12.4-19.9 | 5.6 | 0.58 | 22.9 | 4.5- 7.4 | 0.21 | 0.05 | 52.9 | 0.04-0.32 |
| 7 | 13.2 | 1.4 | 18.1 | 10.5-15.0 | 5.0 | 0.70 | 24.2 | 3.9- 6.3 | 0.25 | 0.08 | 53.9 | 0.10-0.36 |
| Overall mean | 15.8 | 0.8 | 27.7 | 6.4-26.7 | 5.8 | 0.33 | 31.4 | 2.4-10.8 | 0.15 | 0.01 | 59.5 | 0.04-0.49 |

Table 13. 1987 phenotypic and life-history traits (continued)

| | Character[a] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 26. Number of Root Nodules (NRN) | | | | 27. Number of Pods per Plant (PPP) | | | | 28a. Proportion of One-Seed Pods (P1SP) | | | |
| Population[b] | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 90.4 | 15.8 | 39.2 | 56-133 | 1208.1 | 66.9 | 17.5 | 920-1504 | 0.16 | 0.04 | 79.9 | 0.06-0.48 |
| 2 | 77.2 | 25.2 | 43.9 | 51-135 | 974.2 | 28.8 | 9.4 | 827-1111 | 0.09 | 0.01 | 34.2 | 0.06-0.16 |
| 3 | 88.2 | 24.2 | 54.8 | 27-127 | 925.4 | 55.8 | 19.1 | 569-1210 | 0.24 | 0.02 | 23.5 | 0.18-0.33 |
| 4 | 74.5 | 11.9 | 31.8 | 53-101 | 806.3 | 29.3 | 10.9 | 684- 935 | 0.19 | 0.01 | 19.8 | 0.13-0.23 |
| 5 | 71.2 | 6.5 | 20.4 | 55- 88 | 1021.6 | 61.1 | 18.9 | 775-1426 | 0.17 | 0.04 | 72.8 | 0.07-0.49 |
| 6 | 61.2 | 12.8 | 46.8 | 17- 97 | 993.6 | 65.9 | 19.9 | 776-1456 | 0.09 | 0.01 | 31.9 | 0.05-0.12 |
| 7 | 80.7 | 15.4 | 33.1 | 56-109 | 924.3 | 77.4 | 14.5 | 771-1020 | 0.12 | 0.05 | 65.4 | 0.04-0.21 |
| Overall mean | 77.2 | 5.4 | 38.7 | 17-135 | 987.9 | 25.3 | 20.0 | 569-1504 | 0.15 | 0.01 | 59.5 | 0.04-0.49 |

Table 13. 1987 phenotypic and life-history traits (continued)

| | Character[a] | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 28b. Proportion of Two-Seed Pods (P2SP) | | | | 28c. Proportion of Three-Seed Pods (P3SP) | | | | 28d. Proportion of Four-Seed Pods (P4SP) | | | |
| Population[b] | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 0.42 | 0.02 | 12.9 | 0.33-0.49 | 0.42 | 0.04 | 32.4 | 0.13-0.59 | 0.008 | 0.001 | 76.0 | 0.001-0.02 |
| 2 | 0.27 | 0.02 | 18.2 | 0.22-0.38 | 0.61 | 0.02 | 10.1 | 0.48-0.68 | 0.03 | 0.01 | 51.9 | 0.01 -0.06 |
| 3 | 0.37 | 0.01 | 7.2 | 0.31-0.40 | 0.39 | 0.02 | 14.1 | 0.29-0.45 | 0.009 | 0.001 | 41.2 | 0.002-0.01 |
| 4 | 0.35 | 0.01 | 12.1 | 0.27-0.41 | 0.43 | 0.02 | 11.8 | 0.39-0.54 | 0.03 | 0.004 | 39.3 | 0.02 -0.05 |
| 5 | 0.31 | 0.03 | 28.0 | 0.22-0.46 | 0.50 | 0.05 | 33.4 | 0.15-0.66 | 0.02 | 0.005 | 72.8 | 0.001-0.05 |
| 6 | 0.28 | 0.03 | 31.4 | 0.20-0.46 | 0.60 | 0.03 | 14.8 | 0.41-0.61 | 0.03 | 0.01 | 66.2 | 0.001-0.06 |
| 7 | 0.26 | 0.05 | 31.9 | 0.18-0.34 | 0.59 | 0.09 | 26.1 | 0.44-0.74 | 0.03 | 0.01 | 41.9 | 0.01 -0.04 |
| Overall mean | 0.33 | 0.01 | 24.3 | 0.18-0.49 | 0.49 | 0.02 | 27.0 | 0.13-0.74 | 0.02 | 0.002 | 75.8 | 0.001-0.06 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a] 29. Seed Weight per Plant (gm) (SWP) | | | | 30. Seed Weight per One-Hundred Seeds (gm) (SWH) | | | | 31. Number of Seed per Plant (SPP) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range | $\bar{X}$ | SE | CV | Range |
| 1 | 47.2 | 2.0 | 13.5 | 38.7-56.3 | 1.92 | 0.06 | 9.8 | 1.73-2.21 | 2472.6 | 102.0 | 13.0 | 2071-2997 |
| 2 | 40.6 | 2.0 | 15.4 | 26.1-47.9 | 1.83 | 0.06 | 10.6 | 1.54-2.11 | 2244.5 | 122.2 | 17.2 | 1344-2610 |
| 3 | 40.7 | 2.4 | 18.4 | 29.4-53.6 | 2.14 | 0.05 | 7.6 | 1.97-2.54 | 1924.3 | 133.6 | 21.9 | 1159-2722 |
| 4 | 42.1 | 1.8 | 12.6 | 36.1-49.5 | 2.49 | 0.03 | 3.4 | 2.36-2.61 | 1692.0 | 68.7 | 12.2 | 1382-1990 |
| 5 | 43.8 | 2.0 | 14.5 | 31.5-53.9 | 2.00 | 0.12 | 18.5 | 1.70-2.84 | 2255.5 | 170.9 | 23.9 | 1526-3132 |
| 6 | 45.3 | 2.0 | 13.3 | 37.9-55.4 | 1.87 | 0.07 | 11.6 | 1.61-2.38 | 2443.9 | 127.0 | 15.6 | 1810-3096 |
| 7 | 40.7 | 2.6 | 11.1 | 35.5-43.5 | 1.82 | 0.08 | 7.9 | 1.66-1.94 | 2258.7 | 231.3 | 17.7 | 1829-2622 |
| Overall mean | 43.1 | 0.83 | 14.9 | 26.1-56.3 | 2.02 | 0.04 | 15.0 | 1.54-2.84 | 2179.8 | 59.0 | 21.1 | 1159-3132 |

Table 13. 1987 phenotypic and life-history traits (continued)

| Population[b] | Character[a]: 32. Average Number of Seeds per Pod (ASP) | | | |
|---|---|---|---|---|
| | $\bar{X}$ | SE | CV | Range |
| 1 | 2.07 | 0.06 | 8.9 | 1.71-2.25 |
| 2 | 2.30 | 0.10 | 14.0 | 1.44-2.54 |
| 3 | 2.07 | 0.03 | 5.0 | 1.89-2.25 |
| 4 | 2.10 | 0.04 | 6.4 | 1.89-2.36 |
| 5 | 2.21 | 0.10 | 14.2 | 1.53-2.64 |
| 6 | 2.48 | 0.05 | 6.5 | 2.13-2.64 |
| 7 | 2.44 | 0.12 | 8.4 | 2.78-2.67 |
| Overall mean | 2.21 | 0.03 | 11.7 | 1.44-2.67 |

[a] For explanation of character notation, see Materials and Methods

[b] Populations: 1= Hatsunedai; 2= Kanodanchi; 3= Nishi-asahigaoka; 4= Yata; 5= Asahigaoka-1; 6= Asahigaoka-2; 7= Kakitagawa.

D. Analysis of variance (ANOVA)

Of the 32 traits examined, 21 were significantly different among G. soja populations in 1986 and 18 were significantly different in 1987 (Table 14). CTL, LCL and LLL were significantly different in 1986, but not in 1987. As the methods of selecting and measuring flowers and leaves were the same for both years, the reasons for the differences in results for these three traits is not intuitively obvious. One possibility is that year-to-year fluctuations in the greenhouse environment (e.g. temperature, lighting, soil conditions) elicited different responses from the wild soybean populations with respect to these flower and leaf traits.

Leaf, flower and certain yield characters had large among-population components of their variances (as indicated by their high $r^2$ values);however, most of the variation in quantitative traits was found within rather than among populations (Table 14). The $r^2$ values ranged from 0.08 to 0.90 in 1986, with an average of 0.32. In 1987, $r^2$ ranged from 0.03 to 0.66, with an average value of 0.29. Thus, about 70% of the total phenotypic variation was due to within-population variation in both years.

E. Canonical discriminant analysis (CDA)

Relationships among the seven populations were explored further using canonical discriminant analysis. Canonical

Table 14. Results of a one-way ANOVA for quantitative traits among greenhouse-grown populations of wild soybean.

| Dependent variable[a] | | Year 1 (1986) | | Year 2 (1987) | |
|---|---|---|---|---|---|
| | | $r^2$[b] | F-ratio | $r^2$ | F-ratio |
| 1. | DTG | 0.08 | 0.75 | 0.03 | 0.31 |
| 2. | DTF | 0.21 | 2.32* | 0.45 | 7.32** |
| 3. | DFP | 0.28 | 3.46** | 0.26 | 3.17** |
| 4. | DDP | 0.10 | 0.99 | 0.19 | 2.12 |
| 5. | LDP | 0.05 | 0.48 | 0.19 | 2.11 |
| 6. | LSN | 0.11 | 1.08 | 0.07 | 0.70 |
| 7. | B4W | 0.15 | 1.58 | 0.14 | 1.46 |
| 8. | L4W | 0.13 | 1.28 | 0.12 | 1.19 |
| 9. | H4W | 0.27 | 3.33** | 0.18 | 1.95 |
| 10. | GH | 0.35 | 4.68** | 0.39 | 5.81** |
| 11. | L7W | 0.14 | 1.40 | 0.10 | 1.01 |
| 12. | ONO | 0.20 | 2.15 | 0.11 | 1.11 |
| 13. | BPW | 0.90 | 78.18** | 0.51 | 9.36** |
| 14. | FL | 0.79 | 32.90** | 0.26 | 3.18** |
| 15. | CTL | 0.31 | 3.87** | 0.12 | 1.17 |
| 16. | LCL | 0.34 | 4.49** | 0.20 | 2.23 |
| 17. | WCL | 0.50 | 8.65** | 0.55 | 11.06** |
| 18. | LRC | 0.51 | 8.83** | 0.66 | 17.86** |
| 19. | LLL | 0.34 | 4.49** | 0.03 | 0.24 |
| 20. | WLL | 0.31 | 3.83** | 0.33 | 4.46** |
| 21. | LRL | 0.46 | 7.51** | 0.52 | 9.90** |
| 22. | RF[c] | ---- | ---- | ---- | ---- |
| 23. | AGW | 0.21 | 1.13 | 0.37 | 2.34 |
| 24. | BGW | 0.10 | 0.48 | 0.25 | 1.33 |
| 25. | NDW | 0.25 | 1.47 | 0.35 | 2.18 |
| 26. | NRN | 0.19 | 1.04 | 0.11 | 0.48 |
| 27. | PPP | 0.40 | 5.95** | 0.36 | 5.06** |
| 28a. | P1SP | 0.32 | 4.19** | 0.32 | 4.28** |
| 28b. | P2SP | 0.47 | 7.93** | 0.43 | 7.34** |
| 28c. | P3SP | 0.45 | 7.23** | 0.49 | 6.93** |
| 28d. | P4SP | 0.58 | 12.24** | 0.39 | 5.82** |
| 29. | SWP | 0.18 | 1.91 | 0.14 | 1.53 |
| 30. | SWH | 0.54 | 10.36** | 0.53 | 10.06** |
| 31. | SPP | 0.34 | 5.08** | 0.34 | 4.75** |
| 32. | ASP | 0.33 | 4.43** | 0.34 | 4.68** |

[a] Plant origin (population) is the independent variable (d.f. = 6).

[b] $r^2$ is the proportion of variance in the dependent variable which is accounted for by the model.

[c] Root fluorescence (RF) was negative for all plants tested.

* and **: Difference significant at the 5% and 1% levels, respectively.

variates were extracted from a 22-member data matrix. The matrices for 1986 and 1987 contained identical variables. Canonical variates represent orthogonal, least-squares lines through multivariate space defined by the characters in the data matrix. The between canonical functions and the standardized canonical coefficients for 1986 and 1987 are given in Appendix VIII.

1. 1986. Three of the canonical variates extracted from the 22-member phenotypic data set in 1986 were significant at the 1% level as indicated by Wilk's lambda. These three variates accounted for 53.1%, 35.4% and 5.3% of the total phenotypic variation, respectively. Variables which contributed most (based on the between canonical structure) to the first canonical variate were LDP, LWL, ONO, FL, CTL, and SWH. Variables which contributed most to the second canonical variate were DTG, DFP, GH, BPW, PPP, SPP and SWP. LLL and LCL contributed most to the third canonical variate.

A scattergram of the class means on the canonical variates for the wild soybean populations relative to the first two canonical variates is shown in Figure 25. Populations 1, 2, 5, and 6 clustered together. Populations 3 and 7 were distinct from this group, but not necessarily from each other. Population 4 was distinct from all other populations.

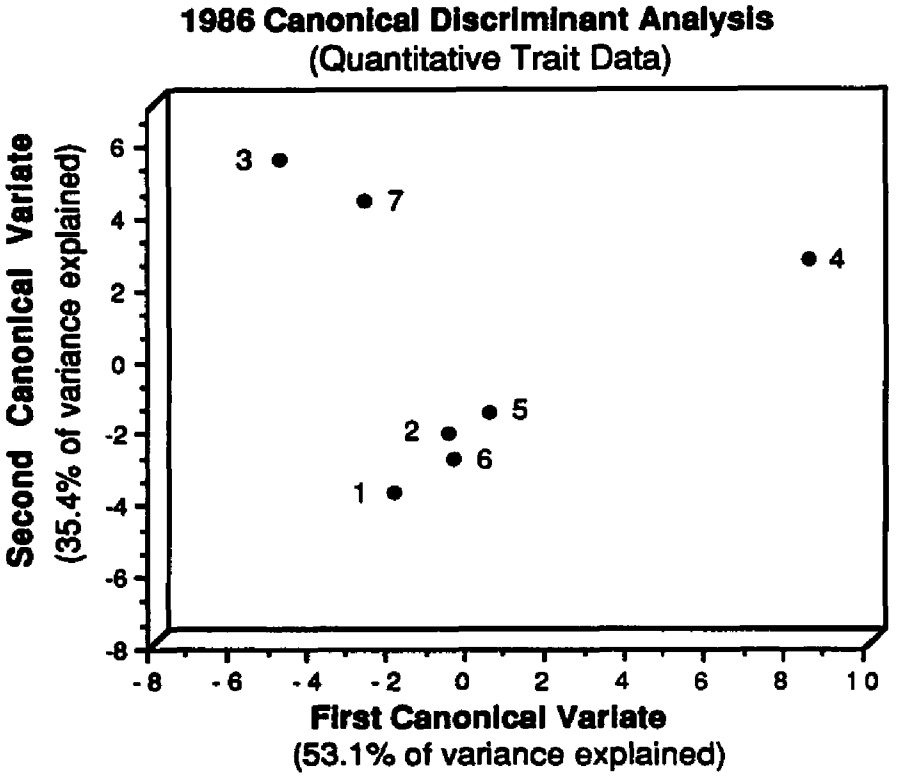



Figure 25. Scattergram of population means on the first two canonical variates. Based on a canonical discriminant analysis of a 22-member quantitative trait data matrix (1986 data). Populations: 1=Hatsunedai; 2=Kanodanchi; 3=Nishi-asahigaoka; 4=Yata; 5=Asahigaoka-1; 6=Asahigaoka-2; 7=Kakitagawa.

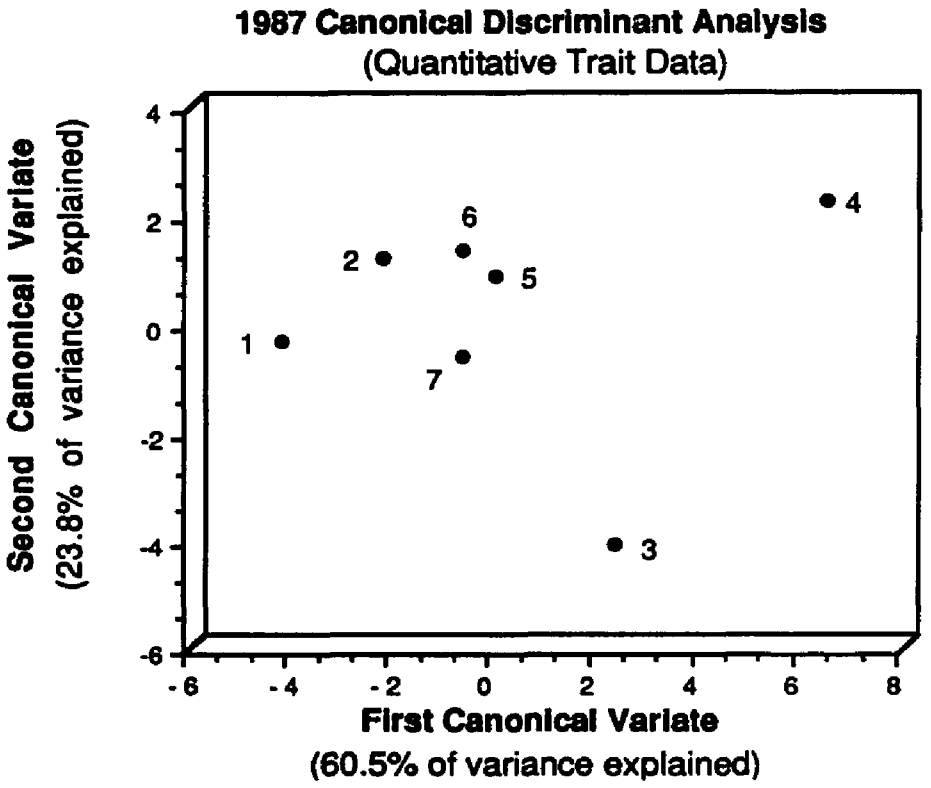



Figure 26. Scattergram of population means on the first two canonical variates. Based on a canonical discriminant analysis of a 22-member quantitative trait data matrix (1987 data). Populations: 1=Hatsunedai; 2=Kanodanchi; 3=Nishi-asahigaoka; 4=Yata; 5=Asahigaoka-1; 6=Asahigaoka-2; 7=Kakitagawa.

2. 1987. Two significant canonical variates (as indicated by Wilk's lambda) were extracted from the phenotypic data matrix in 1987. These variates accounted for 60.5% and 23.8% of the total phenotypic variation, respectively. The characters which contributed most to the first canonical variate were L7W, WCL, PPP, SPP AND SWH. The traits which contributed most to the second canonical variate were BPW, FL and CTL.

A scattergram of the class means on the canonical variates for the seven populations relative to the first two canonical variates is shown in Figure 26. Populations 1, 2, 5, 6 and 7 clustered together. Populations 4 and 3 were distinct from the other populations and from each other.

The results for CDA for the 1986 and 1987 phenotypic data sets are roughly similar. In both years, flower, leaf and yield characters were most effective at discriminating among populations as indicated by their high loadings on the significant canonical variates. One major difference is that population 7 grouped with population 3 in 1986, but with the cluster of populations 1,2, 5 and 6 in 1987 (Figures 25 and 26). Given that population 7 had only three representatives, the instability of group assignment is not surprising. Another difference was the orientation of the clusters in multivariate space. In 1986, the cluster of populations 1, 2, 5 and 6 was in the lower left-hand corner of the scattergram, while populations 3 and 7 were in the upper

left-hand quadrant. In 1987, the orientation of these clusters (except for population 7) was reversed. Population 4 remained in the upper right-hand quadrant of the scattergram in both 1986 and 1987 analyses. The differences in group orientation most likely are due to which characters loaded on the canonical variates in each year.

F. Mahalanobis' distance ($D^2$)

The summaries of phenotypic distance among the populations of wild soybean as indicated by $D^2$ are given in Tables 15 and 16. In both years, population 4 (Yata) had the greatest mean $D^2$. These results parallel those of the isozyme analysis described in Chapter II which demonstrated that Yata had the greatest average genetic distance (D = 0.143) among the seven wild soybean populations. For the 1986 data set, phenotypic distance was greatest between populations 3 (Nishi-asahigaoka) and 4 (Yata) ($D^2$ = 13.56). Populations 2 (Kanodanchi) and 6 (Asahigaoka-2) were most similar phenotypically ($D^2$ = 3.08). For the 1987 data set, the greatest phenotypic distance was between populations 1 (Hatsunedai) and 4 (Yata) ($D^2$ = 11.02). As was the case in 1986, populations 2 (Kanodanchi) and 6 (Asahigaoka-2) were most similar in 1987 ($D^2$ = 2.36).

The difference in which populations had the greatest $D^2$ value in the 1986 and 1987 data sets may have been due to yearly weather-induced fluctuations in the greenhouse

Table 15. Mahalanobis' distance ($D^2$) among seven populations of wild soybean. Distances based on 21 phenotypic and life-history traits. Data collected in the greenhouse in 1986.

| Population | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean $D^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | ____ | 4.43 | 9.85 | 12.36 | 5.23 | 3.82 | 9.70 | 7.56 |
| 2 | | ____ | 9.05 | 10.44 | 3.29 | 3.08 | 7.75 | 6.34 |
| 3 | | | ____ | 13.56 | 9.26 | 9.63 | 5.35 | 9.45 |
| 4 | | | | ____ | 9.48 | 10.61 | 12.03 | 11.41 |
| 5 | | | | | ____ | 4.05 | 7.82 | 6.52 |
| 6 | | | | | | ____ | 8.72 | 6.65 |
| 7 | | | | | | | ____ | 8.56 |

Table 16. Mahalanobis' distance ($D^2$) among seven populations of wild soybean. Distances based on 21 phenotypic and life-history traits. Data collected in the greenhouse in 1987.

| Population | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean $D^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | ____ | 4.37 | 8.03 | 11.02 | 5.13 | 5.61 | 5.16 | 6.55 |
| 2 | | ____ | 7.08 | 9.16 | 2.84 | 2.36 | 4.16 | 4.99 |
| 3 | | | ____ | 7.75 | 5.62 | 6.43 | 5.69 | 6.77 |
| 4 | | | | ____ | 6.93 | 7.90 | 8.33 | 8.52 |
| 5 | | | | | ____ | 2.55 | 3.76 | 4.47 |
| 6 | | | | | | ____ | 4.42 | 4.88 |
| 7 | | | | | | | ____ | 5.25 |

environment that elicited different morphological responses from the wild soybean populations. As discussed previously, a common garden experiment cannot adequately control or account for phenotypic plasticity.

G. Patterns of phenotypic and environmental variation

Evidence for ecotype formation in response to environmental variation was explored using Spearman rank correlations between 1) population means for the quantitative traits and the environmental variables estimated from the soybean seed collection site descriptions (Appendices I and II) and 2) mean Mahalanobis' distance ($D^2$) values and the environmental variables.

Among all the significant correlations detected, only a negative correlation between below-ground weight (BGW) and topography (TOP) was common to 1986 and 1987. There is no apparent underlying biological or ecological reason for this correlation. The exact nature of the topographical differences among the populations is not known. The lack of a general common pattern of correlations between the two data sets is probably the result of heterogeneity in the data sets which is, in turn, due to the relatively small sample sizes used in this study.

The correlation between Mahalanobis' distance ($D^2$) and the percentage of the soybean seed collection site occupied by G. soja relative to other, co-habitating plant species

(COVER) were significant for both 1986 and 1987 data sets ($r_s$=0.87, p=0.01 and $r_s$=0.79, p=0.03, respectively). $D^2$ calculates phenotypic differences among populations by considering all variables (and their relationships to one another) simultaneously (Tabachnik and Fidel, 1983). Thus, $D^2$ summarizes phenotypic differences among populations more effectively than methods which consider each trait individually (Atchley, 1971). This may explain why individual traits did not correlate significantly with COVER while the multivariate phenotypic summary statistic, $D^2$, was significantly correlated.

The strong, positive correlation of $D^2$ with COVER suggests that the suite of traits associated with high $D^2$ values is a population response to the relative percentage of wild soybean plants at a collection site. Plants from areas of dense soybean coverage are more vine-like in their growth habit and tend to produce fewer, but larger and heavier seeds. A vine-like growth habit may be important in dense stands, allowing individuals to climb up and among neighboring plants to reach sunlight (Kiang et al., 1987; Anotonovics et al., 1988). Linhart (1974) observed that plants from dense populations produced larger, heavier seeds than plants from less dense stands. Presumably this k-selected reproductive strategy ensures high survival rates of seeds and seedlings in highly competitive environments, especially where resources are limited (Harper et al.,

1970).

A creeping growth habit may be advantageous in less dense areas, allowing for increased dispersal of progeny at the time of seed dispersal, especially in plants with ballistic seed dispersal, like G. soja. In low density stands, intra-population competition for resources may not be as critical as in dense stands; therefore, an r-selected reproductive strategy, where many light seeds are produced, may increase the likelihood of a seed being deposited in a suitable environment for germination and growth and reproduction (Linhart, 1974).

For the most part, the suites of characters associated with the seven soybean populations fit this model (Table 17). One exception is population 7 which has a moderate $D^2$ value, but a high rating for COVER. As population 7 had only three representatives, the accuracy of its $D^2$ value is questionable. It also is possible the the plants in population 3 were distributed widely across the site rather than in a single, high-density stand.

Table 17. Rank order of wild soybean populations with respect to Mahalanobis' distance ($D^2$) and COVER (the relative abundance of wild soybean plants at a collection site).

| Population | $D^2$ [a] | COVER[b] |
|---|---|---|
| 4 | 9.96 | high |
| 3 | 8.11 | moderate |
| 1 | 7.06 | moderate |
| 7 | 6.91 | high |
| 6 | 5.76 | low |
| 2 | 5.66 | low |
| 5 | 5.50 | low |

[a] $D^2$ averaged over 1986 and 1987.
[b] low = < 30% ; moderate = 31-50% ; high = > 50%

# DISCUSSION

## Summary

Univariate and multivariate statistical analyses revealed significant genetic differences among the seven natural populations of G. soja for many of the quantitative characters examined. A univariate analysis of variance indicated that the wild soybean populations retain a large proportion (approximately 70%) of their phenotypic variation (and, therefore, evolutionary potential) within populations. This pattern of population structure for quantitative traits is common among natural populations of predominantly self-fertilizing plants (Allard et al., 1968; Giles and Edwards, 1983; Clay and Levin, 1989). Canonical discriminant analysis demonstrated that the populations of wild soybean have diverged most significantly for those characters related to flower size, leaf shape, and yield.

## Phenotypic and environmental variation in wild soybean

Summarizing years of research on quantitative variation in inbreeding populations, Allard et al. (1968) concluded,

> Clinal variation is frequently observed in association with progressive changes in rainfall, temperature and other factors of the physical environment. Superimposed on such broad geographical variations is a patchwork or mosaic pattern of variation which reflects adaptations to the local environment. The pattern is often a very fine one and since the numbers of niches included

> in areas as small as a few square meters may be considerable, such local differentiation appears to provide for massive storage of genetic variability.

The results of past and present research on phenotypic and life-history variation in the wild soybean are consistent with this generalization. Using the "common garden" method, Chiang (1985) detected significant differences among 12 accessions and 4 natural populations of G. soja from South Korea and Japan for a number of morphological, agronomic and phenological traits. Using principal components analysis, she demonstrated that much of the phenotypic variation among the accessions and the natural populations could be accounted for by latitude (Chiang, 1985; Kiang and Chiang, 1989). Soybean plants from the northern latitudes initiated flowers earlier, had higher numbers of seeds per pod, had lower rates of ovule and seed abortion and devoted a greater proportion of their life-span to reproduction than did plants from the southern latitudes (Chiang, 1985; Kiang and Chiang, 1989). Kiang and Chiang (1989) concluded that these differences were adaptations to the different lengths of the growing season in the northern and southern latitudes.

In the present study, the groupings of the seven wild soybean populations in multivariate space from the canonical discriminant analyses were remarkably consistent over two years. Individuals from these populations are not, therefore, random assemblages of genotypes, but are

structured in a predictable way. That is, group membership in a cluster can be predicted reliably given enough information on phenotypic and life-history characters. The soybean populations were not separated by any obvious progressive changes in the physical environment. Thus, the phenotypic differences observed among these populations may be due to 1) genetic changes resulting from adaptation of individuals to a particular, localized environmental regime as described by Allard et al. (1968), 2) stochastic processes of genetic drift and founder's effects, or 3) some combination of selective and stochastic processes (Lande, 1977; Rai and Jain, 1982).

The results of the Spearman correlation analysis in the present study suggest that population density is one important component of the local environment contributing to population structure and differentiation among the populations of wild soyean. The importance of intra- and inter- specific plant associations as factors affecting population structure have been demonstrated for populations of Anthoxanthum (Antonovics et al., 1988). Unfortunately, neither the distribution of the soybean plants (i.e., whether the plants were spatially scattered or clumped within each collection site) nor the types and distributions of co-habitating plant species were recorded for all collection sites in the present study. It is not possible, therefore, to quantify the effects of such associations in

this instance.

The wild soybean populations from which individuals were sampled for this study were relatively small (H.I. Oka, personal communication) and gene flow between populations was highly improbable. Thus, it is likely that genetic drift and founder's effect have played some role in conditioning the observed patterns of phenotypic variation. However, without additional information on the natural histories of the wild soybean populations, it is not possible to estimate the relative contributions of selective and stochastic forces (Rai and Jain, 1982).

## Future research directions

The results of the present study point to the need for long-term monitoring of quantitative trait variation in natural populations *G. soja*. The following five areas in particular need to be addressed before a meaningful understanding of population structure is achieved: 1) stability and dispersion of soybean populations over time, 2) rates of gene flow among individuals, 3) changes in gene frequencies over time and space, 4) degree of phenotypic plasticity within and among populations, and 5) determination and measurement of the "selective environment" (i.e., that part of the external environment which directly affects differential contribution of genotypes to subsequent generations) (Antonovics et al., 1988).

# CHAPTER IV

## CONGRUENCE BETWEEN ISOZYMES AND QUANTITATIVE TRAITS IN DESCRIBING POPULATION DIFFERENTIATION

### INTRODUCTION

Isozymes and quantitative traits provide two character sets with which to analyze genetic structure and population differentiation. As multilocus associations are a more common feature of selfing plants than of outcrossers (Brown, 1979), some degree of concordance is expected between the two character sets in describing differentiation among populations of predominantly selfing plants (Price et al., 1984).

Reports of statistically significant associations between isozyme variation and quantitative trait variation in populations of plants is widespread, but not universal. Isozymes have been used successfully as markers of quantitative traits in breeding programs for Lycopersicon (tomato) and Solanum (potato) (Tanksley et al., 1981) and Zea (corn) (Stuber et al., 1982). Graef et al. (1989) found significant associations between enzyme locus genotypes and five agronomic traits among interspecific soybean crosses (G. max X G. soja), but the associations were population specific. In studies of plant population structure, the results also have been mixed. Marshall and Allard (1970)

found that in Avena barbata (slender oat) and A. fatua, "populations which [were] monomorphic for morphological markers also [tended] to be monomorphic for isozyme variants, and vice versa." Hamrick and Allard (1975) also found strong positive associations between quantitative characters and isozymes among populations of A. barbata. However, studies of other populations in this species revealed no evidence of significant associations between the two character sets (Allard et al., 1968; Kahler et al., 1980). Similarly, Nevo et al. (1979) found positive correlations between isozymes and quantitative traits in Hordeum spontaneum (barley), whereas Giles (1984) detected no such associations in Hordeum murinum. Price et al. (1984) examined biochemical and quantitative trait variation in three inbreeding plant species (Avena barbata, Hordeum jubatum and Hordeum vulgare) and in one outcrossing species (Clarkia williamsonii). Among populations of Avena, the congruence between quantitative and biochemical characters was statistically significant; for the two species of Hordeum and the outcrosser,. Clarkia, the degree of association was not significant at the 5% probability level. Chiang (1985) found significant associations between quantitative traits and isozymes among twelve accessions of G. soja from South Korea and Japan, but not among four natural populations of wild soybean from Japan.

It is evident from these studies that association of

biochemical and quantitative traits is a population-specific phenomenon rather than a general rule for inbreeding plant species. Thus, no single class of character can be relied on to give a complete understanding of population structure or evolutionary divergence (Allard et al., 1968). It is important to note that associations between biochemical and quantitative traits are not necessarily the result of a common genetic basis for the two character types. Small population sizes due to sampling effects (i.e., population bottlenecks, founder's events) combined with restricted gene flow due to high rates of inbreeding and/or low rates of migration also can lead to such associations (Bryant, 1984).

Another important aspect to detecting and quantifying the relationship between isozymes and quantitative trait associations is the different natures of these two character types. Quantitative traits, because they are coded for by many gene loci, represent a larger proportion of the genome than would an equal number of standard electrophoretic variants (Lewontin, 1984). In addition, expression of quantitative traits often is conditioned by additive, dominance, and epistatic effects, as well as phenotypic plasticity in response to environmental variation. Enzyme polymorphisms, on the other hand, usually are discrete markers of individual gene loci whose expression is not generally affected by the environment (Brown and Weir, 1983). Given these differences, it is not surprising that the

association between biochemical and quantitative traits often is low. Lewontin (1984) further warns that the underlying distributions of isozyme variants and quantitative traits may be very different and that current statistical approaches and assumptions often have led to unwarranted conclusions about the rates and nature of evolutionary divergence of these traits.

In this chapter, I will determine the extent of congruence between isozymes and phenotypic traits in describing population differentiation among the seven natural populations of G. soja.

## MATERIALS AND METHODS

### A. Cluster analysis

Cluster analysis was used to produce phenetic "trees" (phenograms) reflecting the relationships among the seven populations of *G. soja* based on the electrophoretic and quantitative trait data sets. Squared Euclidean distance ($d^2$) (Wishart, 1987) was used as the measure of dissimilarity among populations for cluster analysis. The squared Euclidean distance is the sum of squared differences in values for each variable (Norusis,1985). The squared Euclidean distance between two populations, a and b, is defined as

$$d^2_{ab} = 1/n \sum_{i=1}^{n} (X_{ia} - X_{ib})^2 ;$$

where n is the number of traits, and $X_{ia}$ and $X_{ib}$ are the values of the ith character in populations a and b, respectively (Wishart, 1987). Electrophoretic and quantitative data were standardized to z-scores (Sokal and Rohlf, 1969; Wishart, 1987) prior to clustering using the STANDARDISE command in CLUSTAN (Wishart, 1987).

The HIERARCHY subroutine in the CLUSTAN 3.2 computer package (Wishart, 1987) was used to perform the analyses. The clustering algorithm of Ward (Ward, 1963) was chosen. Ward's method is hierarchical and agglomerative. It creates mutually

exclusive groups of entities that are maximally similar with respect to a set of characteristics, minimizing the loss of information associated with the process of clustering. The resulting phenograms do not necessarily reflect phylogenetic relationships among the groups.

Allele frequencies were used as the data matrix for the electrophoretic phenogram. Population means for the 22 quantitative characters used in canonical discriminant analysis in Chapter III were used as the data matrices for the 1986 and 1987 quantitative trait phenograms (See Appendix VIII).

B. Congruence between isozymes and quantitative trait data

The congruence between the isozyme and quantitative trait data sets in describing differentiation among the wild soybean populations was explored using three methods. First, the phenograms generated from cluster analysis were compared for the congruence of their branching topologies using Farris' cluster distortion method (Farris, 1973). The cluster distortion coefficients may range from 0 to 1, with a value of 0 indicating complete congruence and a value of 1 indicating complete distortion. Second, the degree of association between Nei's genetic distance (D, from the electrophoretic data), Mahalanobis' distance ($D^2$, from the canonical discriminant analysis using quantitative traits) and squared Euclidean distance ($d^2$, from the cluster analysis

using the quantitative traits) was tested using Spearman rank correlations. Third, the average coefficient of variation (CV) within each population (pooled over all characters, see Tables 12 and 13 in Chapter III) and the number of alleles per isozyme locus (A) were compared using a Spearman rank correlation. As the number of alleles per isozyme locus is sensitive to sampling error, a significant correlation between the average CV and A would indicate a possible role for founder's events or bottlenecks in shaping population structure (Bryant, 1984).

C. Spearman rank correlations between isozyme loci and quantitative traits

Because of the non-normal distribution of allele frequencies, the relationships between enzyme loci and quantitative traits were explored using Spearman rank coefficients. Correlations between quantitative traits were tested using Pearson product-moment correlation analysis because the distributions of these traits were known to be approximately normal (see Chapter III). The detection of significant and consistent associations of specific enzyme loci with specific quantitative traits may provide useful markers for inter-specific soybean breeding programs using G. soja germplasm (Kiang and Gorman, 1983; Graef et al., 1989) and aid efforts toward mapping isozyme and morphological traits on the soybean genome.

## RESULTS

### A. Cluster analysis

Phenograms of the seven populations of wild soybean based on electrophoretic and quantitative trait data are shown in Figure 27. The dissimilarity matrices for the three data sets, from which the phenograms were derived, are given in Appendix X. The electrophoretic and 1987 quantitative trait phenograms were very similar in their branching topologies, having essentially three population clusters. The first cluster consisted of populations 3 and 4; the second consisted of populations 2, 5, 6 and 7 ; the third contained population 1 only. The 1986 phenotypic trait phenogram was very different from the other two phenograms. In the 1986 tree, population 4 was in a cluster by itself; the second cluster consisted of populations 2, 3, 6 and 7; the third cluster contained populations 1 and 5.

The groupings of populations using cluster analysis differed somewhat from the groupings obtained through canonical discriminant analysis (see Chapters II and III). For example, in canonical discriminant analysis of the phenotypic trait data, populations 3 and 4 were distinct from one another in multivariate space. However, these two populations were grouped together in two of the cluster analyses. The differences are most likely due to the very

different natures of these analyses. CDA considers all variables (and their relationships to one another) simultaneously to find the most parsimonious suite of characters which best differentiates groups in multivariate space. Cluster analysis, on the other hand, groups entitities which are most similar based on squared Euclidean distance.

Farris' cluster distortion coefficients, which indicate the degree of congruence between the branching topologies of any given pair of trees, are given in Table 18. The congruence between the electrophoretic and 1986 quantitative trait phenogram was low (distortion=0.70). Conversely, the congruence between the electrophoretic and 1987 quantitative trait phenogram was moderate (distortion=0.20). Between the 1986 and 1987 quantitative trait phenograms, the congruence was 0.66.

B. Comparison of D, $D^2$, and $d^2$

The Spearman rank correlation analysis between Nei's (D), Mahalanobis' ($D^2$) and squared Euclidean ($d^2$) distance gave results similar to those of the cluster analysis (Table 19). For 1986, the degree of association between isozymes and quantitative traits was relatively high, but not statistically significant. The correlation between D and $D^2$ was 0.61 ($p=0.14$); the correlation between D and $d^2$ was 0.67 ($p=0.10$). For 1987, the degree of association was statistically significant. The correlation coefficient

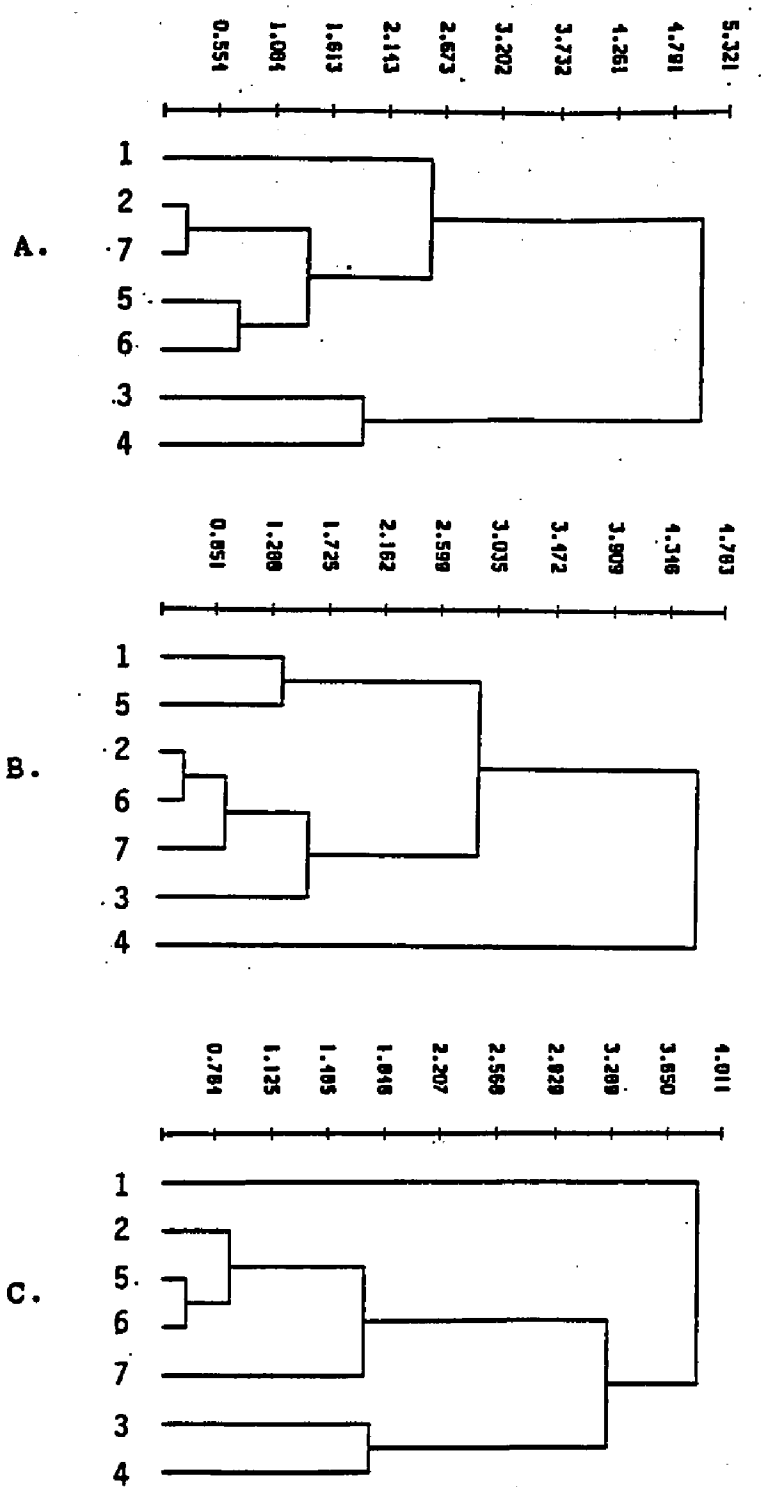



Figure 27. Ward's clustering of seven populations of G. soja using allele frequencies (A) and quantitative traits of plants grown in the greenhouse in 1986 (B) and 1987 (C). Axes represent cluster fusions based on squared Euclidean distances. Populations: 1=Hatsunedai; 2=Kanodanchi; 3=Nishi-asahigaoka; 4=Yata; 5=Asahigaoka-1; 6=Asahigaoka-2; 7=Kakitagawa.

Table 18. Cluster distortion coefficients between electrophoretic and quantitative phenograms of seven natural populations of wild soybean.

| Phenograms compared[a] | Cluster distortion coefficient | Standard deviation |
|---|---|---|
| Electrophoretic (A) versus 1986 Quantitative (B) | 0.70 | 0.04 |
| Electrophoretic (A) versus 1987 Quantitative (C) | 0.20 | 0.08 |
| 1986 Quantitative (B) versus 1987 Quantitative (C) | 0.66 | 0.01 |

[a] See Figure 27.

Table 19. Spearman rank correlation coefficients among Nei's (D), Mahalanobis' ($D^2$) and squared Euclidean ($d^2$) distance. Results for 1986 above the diagonal; results for 1987 below the diagonal.

| Distance measure | Distance measure | | |
|---|---|---|---|
| | D | $D^2$ | $d^2$ |
| D | ____ | 0.61 | 0.67 |
| $D^2$ | 0.85* | ____ | 0.68 |
| $d^2$ | 0.77* | 0.86* | ____ |

*: correlation coefficient significantly different from zero at the 5% level.

between D and $D^2$ was 0.85 (p=0.02); the correlation between D and $d^2$ was 0.77 (p=0.04).

C. Comparison of mean CV and A

The mean coefficient of variation (CV) for each population in 1986 and 1987 and the average number of alleles per isozyme locus (A) are given in Table 20. The Spearman rank correlation between CV and A was not significant in 1986 ($r_s$=0.34, p=0.45), but was significant in 1987 ($r_s$=0.78, p=0.04). Thus, it is possible that sampling error due to founder's events or population bottlenecks have played a role in shaping the genetic structure of the populations of wild soybean examined here, but the evidence is not conclusive.

D. Spearman rank correlations between isozyme loci and quantitative traits

Among all significant correlations detected between enzyme loci and quantitative traits, 19 were consistent across both years of the study. These correlations are summarized in Table 21. The complete listing of correlation coefficients is given in Appendix XI.

Among the 19 correlations listed in Table 21, two patterns were particularly interesting. The characters, days to last dry pod (LDP), leaflet shape (LRC and LRL) and the number of pods per plant (PPP), were consistent as a block of traits significantly correlated with four of the enzyme locus genotypes (Aco2, Am3, Dia2, Lap). The four phenotypic

Table 20. Spearman rank correlation between average coefficient of variation (CV) for quantitative traits and mean number of alleles per isoyzme locus (A) among seven natural populations of wild soybean.

| Population | Mean CV | | Mean number of alleles per locus |
|---|---|---|---|
| | 1986 | 1987 | A |
| 1 | 19.6 | 21.4 | 1.19 |
| 2 | 20.2 | 16.8 | 1.19 |
| 3 | 21.7 | 22.3 | 1.03 |
| 4 | 14.1 | 11.8 | 1.00 |
| 5 | 21.3 | 23.8 | 1.25 |
| 6 | 25.9 | 23.4 | 1.23 |
| 7 | 24.8 | 20.2 | 1.08 |
| Spearman rank correlation between CV and A | 0.34 | 0.78* | |

*: correlation significantly different from zero at the 5% level.

Table 21. Spearman rank correlations between isozyme loci and quantitative traits which were consistently significant over a two year study of the wild soybean.

| | Character[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Locus | LDP | GH | WCL | LRC | LRL | PPP | P1SP | P2SP |
| Aco2-a | | | | **(-) | **(-) | *(-) | | |
| Aco2-b | | | | **(+) | **(+) | *(+) | | |
| Am3-s | *(-) | | | *(-) | *(-) | *(-) | | |
| Am3-f | *(+) | | | *(+) | *(+) | *(+) | | |
| Dia1-a | | | *(-) | | | | | |
| Dia1-b | | | *(+) | | | | | |
| Dia2-a | *(-) | | | *(-) | *(-) | *(-) | | |
| Dia2-b | *(+) | | | *(+) | *(+) | *(+) | | |
| Est-a | | **(+) | | | | | **(-) | **(+) |
| Est-b | | **(-) | | | | | **(+) | **(-) |
| Gpd-a | | **(+) | | | | | | |
| Gpd-b | | **(-) | | | | | | |
| Lap-a | | | | *(-) | *(-) | *(-) | | |
| Lap-b | | | | *(+) | *(+) | *(+) | | |

[a] For explanation of character notation, see Chapter III, Materials and Methods.

* and **: correlation coefficient was significantly different from zero at the 5% and 1% level of significance, respectively.

(-) and (+): negative and positive correlations, respectively.

characters were, for the most part, significantly and positively correlated with each other (Table 22), suggesting that late-maturing soybean plants with narrow leaflets yielded more pods per plant than did early-maturing plants with oval leaflets. These results are consistent with the results of an inheritance study in G. soja by Domingo (1945), who observed that narrow-leaflet genes were linked closely with genes for high yield. Another interesting correlation was the association of the esterase locus with early growth habit in soybean. Darmency and Gasquez (1983) detected a similar association in Poa annua, an annual grass.

Whether the correlations between enzyme locus genotypes and quantitative characters detected here reflect functional and/or chromosomal relationships is beyond the scope of the present research project. However, the consistency of the correlations over the course of this study indicates that they merit further investigation. Genetic, biochemical and physiological analyses are needed to assess the relationships among specific enzyme loci and quantitative traits before their potential in breeding programs and mapping studies can be determined.

Table 22. Pearson product-moment correlation coefficients between selected quantitative traits across a two year study of the wild soybean. 1986 above the diagonal; 1987 below the diagonal.

| | Character[a] | | | |
|---|---|---|---|---|
| Character | LDP | LRC | LRL | PPP |
| LDP | ---- | 0.86** | 0.70 | 0.93** |
| LRC | 0.91** | ____ | 0.80* | 0.97** |
| LRL | 0.95** | 0.92** | ____ | 0.76* |
| PPP | 0.93** | 0.92** | 0.94** | ____ |

* and **: correlation coefficient significantly different from zero at the 5% and 1% levels, respectively.

## DISCUSSION

The results of this study indicate, although not conclusively, that there is some degree of concordance between isozymes and quantitative traits among the seven populations of wild soybean examined. To understand the results, two questions must be addressed. First, what factor or factors can be invoked to explain concordance between isozymes and quantitative traits? Second, why was the congruence between isozymes and quantitative traits not consistent across both years of the present study?

Congruence between isozymes and quantitative traits may be due to 1) functional relationships between isozyme variants and quantitative traits, 2) genetic (chromosomal) linkage between isozyme loci and quantitative trait loci, 3) co-adaptation, or 4) chance association due to small population sizes associated with sampling effects (i.e. founder's events, population bottlenecks) combined with predominant inbreeding (Marshall and Allard, 1970; Lande, 1977; Bryant, 1984). Evidence for or against the first three possibilities would require biochemical, physiological and genetic linkage studies far beyond the scope of the present study. Evidence for linkage between some of the enzyme variants and quantitative trait loci was indicated by the Spearman rank correlation analysis, but this evidence is preliminary.

The role of sampling effects in shaping the association between biochemical and quantitative traits was indicated by the significant correlation between the average coefficient of variation (CV) and the average number of alleles per isozyme locus (A) in 1987; however, this correlation was not significant in 1986. Because the number of plants sampled from each population in the field was relatively small, it is possible that the values of A were underestimated and were not representative of the actual amounts of genetic diversity present in the field. Therefore, the significant correlation detected in 1987 may be a spurious one. The other possibility is that the result for 1986 was due to severe heterogeneity in the data sets (due to the small sample sizes) which resulted in a Type II error (accepting a false null hypothesis).

One puzzling aspect to this research was that the congruence between isozymes and quantitative traits was not consistent across both years of the study. The association between biochemical and quantitative traits was non-significant in 1986, but was highly significant in 1987. Three possible explanations for this inconsistency come to mind. First, if functional or genetic linkage relationships do exist between the isozymes and quantitative traits, these relationships may be relatively weak and, therefore, require larger sample sizes than those used here to quantify them accurately and consistently. A second possibility is

experimental error. In 1986, I employed two different technical assistants in the greenhouse to aid in collection of measurement data; in 1987, I performed all measurements myself and employed an assistant only in the collection of dry pods during the harvest. Given the relatively small sample sizes used in this study, even minor differences in measurement technique could significantly affect the results. The third possibility has to do with the nature of the quantitative traits. I do not know how many loci governed the quantitative traits, the extent to which those loci were heterozygous and segregating within populations, or the extent to which quantitative trait variation was due to plastic responses. It is possible that year-to-year differences in temperature, lighting and soil conditions in the greenhouse elicited different morphological responses from the wild soybean plants. The effect of such confounding could have been severe given the small sample sizes used in this study.

Among natural populations of Xanthium strumarium (a monoecious annual) and Bromus rubens (an annual grass), phenotypic plasticity was observed to be a larger component of total phenotypic variation than genetic variance, suggesting that plasticity is the predominant mode of adaptation to environmental variation in these colonizing species (Wu and Jain, 1978; Jain, 1978; Moran et al., 1981). G. soja possesses many of the phenotypic characteristics of

weedy, colonizing plants. Thus, it is likely that phenotypic plasticity plays some role in conditioning the capacity of wild soybeans to grow and reproduce in such diverse habitats (Kiang, personal communication).

Chiang (1985) examined isozyme and quantitative trait variation in four natural populations of wild soybean from the northern Honshu Island in Japan and did not detect any significant associations between the two character sets. She concluded that electrophoretic variants represented a much smaller proportion of the soybean genome than did the quantitative traits and, therefore, the chances for detecting associations between the two character types was very small. This may be a factor in the present study as well, but I believe that additional factors, such as phenotypic plasticity, also need to be considered.

The apparent small effective mating size of the wild soybean populations, the limited potential for gene flow, and the predominantly selfing mating system are characteristics of populations in which I would expect to detect significant association between electrophoretic and quantitative traits (Lande, 1977; Bryant, 1984; Clay and Levin, 1989). Although there were inconsistencies in the statistical significance of such an association in the present study, the general trend supports this expectation. Clearly, additional evidence is needed to verify these conclusions. Future studies, with emphasis on partitioning quantitative variation into genetic,

environmental and plastic components and with detailed information on the natural histories of the study populations, will greatly enhance our understanding of the processes which contribute to population structure in *G. soja*.

# CHAPTER V

# PLASTICITY IN REPRODUCTIVE YIELD COMPONENTS WITHIN AND AMONG WILD SOYBEAN POPULATIONS

## INTRODUCTION

### Plasticity in yield components

Plasticity in reproductive yield characters (e.g., fruits/plant, ovules/flower, seeds/ovule, seed weight) has both genetic and ecological components and may have a profound impact on the fitness and evolution of populations (Bradshaw, 1965). Significant variation, within and among populations, in yield components has been observed in plants in response to photoperiod (Cook, 1975), temperature (McWilliams et al., 1968), moisture (Schimpf, 1977), life-histories (Marshall et al., 1985) and habitat (Thompson, 1981; Lotz and Spoormakers, 1988).

Significant negative correlations among seed yield components have been observed in cultivated crops and are major factors limiting yield (Adams, 1967). Negative correlations among traits related to fitness in natural populations of plants may be important in limiting responses to natural selection and in conditioning directions of life-history evolution (Lande, 1982). Primack (1978) hypothesized that natural populations of plants should show more variability among seed yield components than crop species. He reasoned that natural populations of plants face greater

spatial and temporal heterogeneity and, therefore, might retain a great range of flexibility to regulate seed production at many different points along a developmental sequence.

Some components of yield are more plastic than others. Fruit number, flowers/inflorescence and fruits/flower are predicted to be the most variable while mean seed weight/plant often is the least variable component of yield (Harper et al., 1970). High flexibility of inflorescence number may reflect past selection for this character to respond to resource availability. Low phenotypic plasticity in mean seed weight/plant and seed number/fruit may reflect past selection for a particular seed size and dispersal mechanism (Primack and Antonovics, 1981). In addition, high seed weight is critical to seedling success in many species of plants and may be maintained under strong selective constraints (Bradshaw, 1965; Harper et al., 1970; But see Silvertown, 1989).

## Soybean reproductive bilogy

Components of yield have been characterized in detail for the cultivated soybean, Glycine max, in large part to develop selection methods to increase seed yield (Van Schaik and Probst, 1958a,b). Most of these strategies have been aimed at improving the retention of pods (Weil and Ohlrogge, 1976; Harmond and Brun, 1971). Among soybean cultivars there is

"tremendous unrealized reproductive potential" due to the abortion of reproductive units (Hansen and Shibles, 1978). Total flower and pod abortion rates for indeterminate and determinate soybean cultivars across many maturity groups range from 20 to 80% (Van Schaik and Probst, 1958a,b; Dominguez and Hume, 1978; Hansen and Shibles, 1978; Weibold et al., 1981). The most active flower shedding period in G. max is 1 to 7 days after flowering (Carlson and Lersten, 1987). The highest rate of pod loss occurs when pods are less than 2 cm long (Weibold et al., 1981).

Environmental factors, such as extremes in temperature and humidity (Van Schaik and Probst, 1958b) and water and nutrient deficiencies (Brevedan et al., 1978), have been shown to influence the rates of flower and pod abortion in the soybean. Such factors as the lack of viable pollen (Van Schaik and Probst, 1958b) and lack of fertilization (Abernethy et al., 1977) do not appear to be major contributing factors to the loss of reproductive potential under normal circumstances. Although the environmental component to reproductive loss is large, a considerable genetic component does exist (Van Schaik and Probst, 1985a; Weibold et al., 1981).

## MATERIALS AND METHODS

### A. Planting design and growth conditions

One seed per pot, 3 plants per population were grown in a completely randomized design in the greenhouse in 1987. Only one seed per parent plant was used due to the limited number of seeds collected from the field. If a seed did not germinate, it was eliminated from the analysis. Seed and pot preparation and growth conditions were identical to those described in Chapter III.

### B. Census procedures

Because wild soybeans produce large numbers of inflorescences, a random subset of floral clusters per plant was examined. Three random inflorescences were tagged every other day throughout the flowering period of the plant. Only inflorescences with unopened buds were tagged. The fate of each flower bud, flower, pod and seed of each tagged inflorescence was recorded. Observations were made daily (usually in the morning hours) with few exceptions. The following characters were examined:

1) Number of flower buds per inflorescence (BPI),
2) Number of flowers per inflorescence (FPI),
3) Number of immature pods per inflorescence (IPI),
4) Number of mature pods per inflorescence (MPI),
5) Number of mature seeds per inflorescence (SPI),
6) Weight of mature seeds per inflorescence (mg) (SWI).

Stages of reproductive development arbitrarily were

defined based on detailed studies of reproductive development of G. max (Fehr et al., 1971; Carlson and Lersten, 1987). Comparable developmental studies for G. soja currently are unavailable. If an unopened bud collapsed 1-2 days after an inflorescence was tagged, the event was considered "bud" abortion, although it was possible that pollination had occurred. The early flower stage was defined as the first four days after anthesis. Late flower stage was defined as day 5 (post-anthesis) to the first sign of pod protrusion. Early pods were considered those less than 2 cm in length (Weibold et al., 1981). Late pods were defined as green fruits greater than 2 cm in length. Pods were considered mature when they had dried and turned dark brown or black.

All surviving, mature pods were harvested (by inflorescence) and stored in separate, labelled envelopes. The pods then were examined individually for seed number and weight and for evidence of seed abortion. Seed abortion was indicated by the presence of a small sac-like structure (indicating abortion of a seed very early in its development) or an extremely small black seed. It is possible that some of the reproductive abscission events classified as seed loss actually were abortion of ovules. Thus, the percentages of seed loss calculated for the seven soybean populations may be overestimates. However, most ovule abortion in soybean occurs just after fertilization, in the very early stages of embryo development as the ovule begins to develop into a seed (Aung

et al., 1986). I did not, therefore, expect to see any developed tissue in most instances of ovule abortion.

C. Calculation of rates of reproductive abscission

The inflorescence data for the three plants from each population were combined to estimate rates of reproductive abscission for each wild soybean population.

1. Flower bud abortion. The percentage of buds aborted was calculated as

$$((B-F)/B) \times 100;$$

where B was the number of buds examined per population, and F was the resulting number of flowers.

2. Total flower abortion. The percentage of flowers aborted was calculated as

$$((F-IP)/F) \times 100;$$

where F was the number of flowers (from equation 1), and IP was the resulting number of immature pods.

3. Total pod abortion. The percentage of pods aborted was calculated as

$$((IP-MP)/IP) \times 100;$$

were IP was the number of immature pods (from equation 2), and MP was the resulting number of mature pods.

4. Early and late stage flower and pod abortion. Total flower and pod abortion were further broken down into early and late abortion stages. The percentage of abortion for early flowers was calculated as

((TFP-EFA)/TFP) X 100;

where TFP (total number of flowers possible) was the number of flowers in the poulation which grew into immature pods plus those flowers which ultimately aborted, and EFA was the number of early flower abortion events. The percentage of early pod abortion was calculated in a similar manner.

The percentage of abortion for late flowers was calculated as

((TFP-LFA)/TFP X 100;

where TFP was the number of flowers in the population which grew into immature pods plus those flowers which ultimately aborted, and LFA was the number of late flower abortion events. The percentage of late pod abortion was calculated in a similar manner.

5. Flower and pod abortion combined. The combined percentage abortion for flowers and pods was calculated as

((F-MP)/F) X 100;

where F was the number of flowers per population, and MP was the resulting number of mature pods.

6. Seed abortion. The percentage of seed abortion was calculated as

((TSP-S)/TSP) X 100;

where TSP (total number of seeds possible) was the number of mature seeds per population plus those seeds which ultimately aborted, and S was the number of mature seeds.

D. Statistical analysis

1. Analysis of variance. Reproductive character data over all inflorescences were averaged to give a mean value for each plant. The means for the individual plants then were analyzed using a one-way analysis of variance (ANOVA) for unbalanced data sets (PROC GLM) (SAS Institute, 1985b) to test for differences among the seven populations.

The assumptions of linearity, normality and equality of variances for each reproductive character were checked using the methods described in Chapter III, Materials and Methods (page 110).

2. Coefficients of variation (CV). Coefficients of variation (CV's) are used to compare the amount of variation in populations having different means (Lewontin, 1966). The CV for a particular trait is calculated as

$$CV = s(100)/Y;$$

where s is the standard deviation and Y is the mean of the trait.

The coefficients of variation for the yield characters were calculated to examine 1) the relative variability among individual yield components and 2) the relative variability among populations with respect to the yield characters examined.

3. Correlation analyses. Pearson product-moment coefficients were used to examined correlations among yield components (individual means). Spearman rank correlations were used to explore the association between yield components (population means) and both environmental variables (Appendix II) and genetic diversity estimates based on isozyme data (Chapter II). (Spearman rank was used because the distributions of the environmental and genetic variables were non-Gaussian.) All correlation coefficients were generated using the PROC CORR subroutine in SAS (SAS Institute, 1985a).

## RESULTS

### A. Missing data

Germination for this study was 100%.

### B. Data transformations

Visual examination of computer-generated histograms showing the distribution of each reproductive character over all seven populations revealed significant positive skewness among the characters. Data were log (x + 1) transformed to improve normality (Tabachnick and Fidel, 1983). The value, 1, was added to each observation (x) because the data set contained a number of values equal to 1. The log (base 10) of 1 is zero; thus, without the addition of one to each observation, the transformed data set would have contained a large number of zero values, complicating statistical analyses (Little and Hills, 1978).

### C. Analysis of variance

The population means for the six characters (untransformed data) are given in Table 23. The means over all seven populations were as follows: flower buds per inflorescence (BPI), 3.15; flowers per inflorescence (FPI), 3.03; immature pods per inflorescence (IPI), 2.15; mature pods per inflorescence (MPI), 1.23; number of mature seeds per inflorescences (SPI), 3.01; and, mature seed weight per

Table 23. Means and standard errors (S.E.) for yield components (untransformed data) among seven natural populations of wild soybean. N is the total number of inflorescences observed over three plants from each population.

| Yield component | | Population[a] 1 | 2 | 3 | 4 | 5 | 6 | 7 | Overall mean |
|---|---|---|---|---|---|---|---|---|---|
| BPI | Mean | 3.46 | 3.56 | 2.26 | 3.00 | 3.48 | 3.53 | 2.75 | 3.15 |
| | S.E. | 0.16 | 0.17 | 0.09 | 0.20 | 0.18 | 0.18 | 0.13 | |
| | N | 95 | 83 | 68 | 62 | 80 | 72 | 68 | |
| FPI | Mean | 3.26 | 3.38 | 2.25 | 2.94 | 3.34 | 3.36 | 2.71 | 3.03 |
| | S.E. | 0.14 | 0.15 | 0.09 | 0.18 | 0.17 | 0.16 | 0.13 | |
| | N | 95 | 83 | 68 | 62 | 80 | 72 | 68 | |
| IPI | Mean | 2.20 | 2.49 | 1.70 | 1.87 | 2.47 | 2.18 | 2.12 | 2.15 |
| | S.E. | 0.10 | 0.12 | 0.08 | 0.12 | 0.14 | 0.09 | 0.11 | |
| | N | 91 | 81 | 64 | 55 | 76 | 72 | 65 | |
| MPI | Mean | 1.24 | 1.34 | 1.18 | 1.14 | 1.31 | 1.17 | 1.24 | 1.23 |
| | S.E. | 0.07 | 0.08 | 0.06 | 0.06 | 0.09 | 0.06 | 0.06 | |
| | N | 53 | 53 | 49 | 35 | 35 | 53 | 51 | |
| SPI | Mean | 2.88 | 3.51 | 2.69 | 2.66 | 3.40 | 2.92 | 2.98 | 3.10 |
| | S.E. | 0.22 | 0.22 | 0.16 | 0.23 | 0.29 | 0.19 | 0.22 | |
| | N | 52 | 53 | 49 | 35 | 35 | 53 | 51 | |
| SWI (mg) | Mean | 55.75 | 60.98 | 53.49 | 69.86 | 68.51 | 56.17 | 63.65 | 61.20 |
| | S.E. | 4.14 | 4.13 | 3.82 | 6.14 | 6.02 | 3.92 | 4.75 | |
| | N | 52 | 53 | 49 | 35 | 35 | 53 | 51 | |

[a] Populations: 1=Hatsunedai; 2=Kanodanchi; 3=Nishi-asahigaoka; 4=Yata; 5=Asahigaoka-1; 6=Asahigaoka-2; 7=Kakitagawa.

inflorescence (mg) (SWI), 61.20.

Analysis of variance revealed significant differences among the seven wild soybean populations for BPI, FPI and IPI; no significant differences were detected among the populations for MPI, SPI or SWI (Table 24).

D. Coefficients of variation

In general, the coefficients of variation for all traits examined were similar over all seven populations of wild soybean (Table 25). The CV's (as %) averaged over all populations were, from largest to smallest: seed number per inflorescence (SPI) (26.6) > immature pods per inflorescence (IPI) (24.9) > flower buds per inflorescence (BPI) (23.1) > mature pods per inflorescence (MPI) (22.7) > flowers per inflorescence (FPI) (22.6) > seed weight per inflorescence (SWI) (12.7). The results support the generalization that mean seed weight is one of the least variable components of yield (Harper et al., 1970).

E. Correlations among yield components

Pearson correlation coefficients among individual means for the six yield components are given in Table 26. The lack of significant negative correlations among yield components may be due to the highly plastic nature of the characters under study (Primack, 1978). Also, the indeterminate growth habit of the wild soybean allows adjustment of reproductive

Table 24. Results of a one-way ANOVA for six yield components[a] among seven natural populations of wild soybean.

| Dependent variable[b] | d.f. | $r^2$ [c] | F-value |
|---|---|---|---|
| BPI | 6 | 0.595 | 3.43 * |
| FPI | 6 | 0.570 | 3.09 * |
| IPI | 6 | 0.598 | 3.48 * |
| MPI | 6 | 0.114 | 0.26 |
| SPI | 6 | 0.108 | 0.19 |
| SWI | 6 | 0.092 | 0.34 |

[a] Log (x +1) transformed data.
[b] Plant origin (population) is the independent variable.
[c] $r^2$ is the proportion of variance in the dependent variable which can be accounted for by the model.
*: Significant at the 5% level; others not significant.

Table 25. Coefficients of variation for six reproductive yield components among seven natural populations of wild soybean.

| Character[a] | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Overall mean |
|---|---|---|---|---|---|---|---|---|
| BPI | 23.3 | 23.4 | 18.7 | 28.4 | 24.7 | 22.4 | 20.9 | 23.1 |
| FPI | 22.2 | 22.9 | 18.6 | 21.7 | 24.7 | 20.9 | 27.4 | 22.6 |
| IPI | 27.1 | 24.7 | 24.7 | 28.4 | 28.1 | 20.0 | 22.0 | 24.9 |
| MPI | 24.6 | 25.6 | 20.7 | 19.1 | 25.0 | 21.6 | 22.0 | 22.7 |
| SPI | 28.8 | 23.2 | 23.9 | 26.6 | 27.0 | 26.9 | 29.9 | 26.6 |
| SWI | 12.8 | 10.7 | 14.4 | 13.1 | 11.9 | 12.3 | 13.9 | 12.7 |

[a] Log (x + 1) transformed data.

Table 26. Pearson product-moment correlation coefficients among individual means for six yield components in the wild soybean, Glycine soja.

| | Yield component[a] | | | | | |
|---|---|---|---|---|---|---|
| Yield component | BPI | FPI | IPI | MPI | SPI | SWI |
| BPI | ____ | 0.99* | 0.57* | 0.01 | 0.13 | -0.02 |
| FPI | | _____ | 0.54* | 0 | 0.15 | 0.02 |
| IPI | | | _____ | 0.08 | 0.17 | -0.15 |
| MPI | | | | _____ | 0.87* | 0.69* |
| SPI | | | | | _____ | 0.78* |
| SWI | | | | | | _____ |

[a] See Materials and Methods for explanation of yield character notation.

* Coefficient is significantly different from zero at the 1% level.

output in response to resource availability (Schaal, 1980; Stephenson, 1984).

The patterns of significant correlations in Table 26 reveal that the yield component characters cluster into two groups: 1) BPI, FPI and IPI and 2) MPI, SPI and SWI. In the first group, BPI was positively correlated with both the number of flowers (FPI) (r=0.99) and the number of immature pods per inflorescence (IPI) (r=0.57) which, in turn, were positively correlated with each other (r=0.54). The transition time between these stages is relatively short; therefore, what happens at one stage is likely to influence strongly what happens at the next. In the second group, mature seed weight (SWI) was positively correlated with both the number of mature pods (MPI) (r=0.69) and the number of seeds per inflorescence (SPI) (r=0.78) which, in turn, were positively correlated with each other (r=0.87). The correlations among these traits likely is due to their interdependency as the final products in the reproductive developmental pathway. In other words, seed weight per inflorescence is strongly linked to the number of seeds per inflorescence which, in turn, is strongly linked to the number of mature fruits produced at an inflorescence. The lack of significant correlations between early and late yield components reflects the resource dependent and flexible nature of these characters (Primack and Antonovics, 1981).

## F. Correlations among yield components, environmental variables, and genetic diversity estimates

The relationships among the environmental variables estimated from the soybean collection site descriptions (Appendices I and II), the genetic diversity estimates based on isozyme data (Chapter II) and the population means for log (x + 1) transformed reproductive yield component data were explored using Spearman rank correlations (Table 27).

### Yield components and reproductive characters

No significant associations were detected between environmental variables and the reproductive yield components (Table 27).

### Yield components and genetic diversity estimates

Several significant (5%) correlations were detected among genetic diversity estimates and reproductive yield components. The number of flowers per inflorescence (FPI) was positively correlated with the average number of alleles per locus (A) ($r_s$=0.83), polymorphism at the 95 and 99% levels (P95, $r_s$=0.78; P99, $r_s$=0.83), and the proportion of heterozygotes observed ($H_{obs}$) ($r_s$=0.79). The number of immature pods per inflorescence (IPI) was positively correlated with A ($r_s$=0.80), P95 ($r_s$=0.77), and P99 ($r_s$=0.80). The number of mature pods per inflorescence (MPI) was postively correlated with A ($r_s$=0.77), P95 ($r_s$=0.75), and P99 ($r_s$=0.77). Finally, the number of seeds per inflorescence

Table 27. Spearman rank correlation coefficients[a] among population means for six yield components, six genetic diversity estimates and four environmental variables.

| Environmental[c] and genetic[d] variables | Yield character[b] | | | | | |
|---|---|---|---|---|---|---|
| | BPI | FPI | IPI | MPI | SPI | SWI |
| TOP | -0.49 | -0.18 | -0.34 | -0.39 | -0.18 | 0.29 |
| SHADE | 0 | -0.49 | -0.50 | -0.29 | -0.10 | 0.16 |
| DIST | -0.26 | 0.16 | 0.16 | 0.16 | -0.16 | 0.77 |
| COVER | -0.75 | -0.26 | -0.26 | -0.77 | -0.26 | 0.32 |
| D | -0.18 | -0.45 | -0.57 | -0.66 | -0.76* | -0.02 |
| A | 0.66 | 0.83* | 0.80* | 0.77* | 0.79* | -0.34 |
| P95 | 0.61 | 0.78* | 0.77* | 0.75* | 0.71 | -0.36 |
| P99 | 0.66 | 0.83* | 0.80* | 0.77* | 0.79* | -0.34 |
| $H_{exp}$ | 0.43 | 0.64 | 0.68 | 0.68 | 0.61 | -0.29 |
| $H_{obs}$ | 0.73 | 0.79* | 0.53 | 0.43 | 0.75 | -0.10 |

[a] Only correlations of interest are shown in this table.
[b] Log (x + 1) transformed data.
[c] Based on collection sites descriptions in Appendices I and II: TOP=topography; SHADE=shading; DIST=degree of disturbance; COVER=relative percentage of soybean coverage at the collection site.
[d] Based on isozyme data described in Chapter II: D=Nei's genetic distance; A=average number of alleles per locus; P95, P99=polymorphism at the 95 and 99% levels, respectively; $H_{exp}$=gene diversity; $H_{obs}$=observed proportion heterozygotes.
* Correlation coefficient significantly different from zero at the 5% level of significance.

(SPI) was positively correlated with A ($r_s$=0.79) and P99 ($r_s$=0.79), and negatively correlated with genetic distance (D) ($r_s$=-0.76).

G. Rates and patterns of abortion in the wild soybean

1. Rates of abortion. The percent abortion of reproductive characters (averaged over all seven populations) were, from largest to smallest: pods (61.4) > flowers (32.3) > seeds (4.5) > flower buds (3.5) (Table 28). One exception was noted for population 2 (Kanodanchi) in which flower bud abortion exceeded seed abortion. This result may be due to chance, or to the way flower bud and seed abortion were defined. Combined rates for flower and pod abortion ranged from 62.1% in population 3 (Nishi-asahigaoka) to 82.8% in population 5 (Asahigaoka-1). The average combined flower and pod loss over all populations was 73.5%. In general, populations 1, 2, 5 and 6 (Hatsunedai, Kanodanchi, Asahigaoka-1 and Asahigaoka-2) had increased rates of abortion for all reproductive units. This trend, however, did not hold in all instances. For example, population 4 (Yata) had rates of flower and pod abscission as high or higher than populations 1, 2, 5 and 6. Among all seven populations, populations 3 (Nishi-asahigaoka) had the lowest, or among the lowest, rates of abortion for all yield characters.

Table 28. Percent abortion of flower buds, flowers, pods and seeds (per inflorescence) among seven natural populations of wild soybean.

| Character[a] | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Overall mean |
|---|---|---|---|---|---|---|---|---|
| Flower buds | 5.8 | 5.3 | 0.6 | 2.2 | 4.0 | 4.7 | 1.6 | 3.5 |
| Early flowers | 10.3 | 5.7 | 2.7 | 1.6 | 2.6 | 5.4 | 5.3 | 4.8 |
| Late flowers | 25.2 | 22.7 | 26.7 | 41.3 | 27.4 | 29.8 | 19.5 | 27.5 |
| Flowers (Total) | 35.5 | 28.4 | 29.4 | 42.9 | 30.0 | 35.2 | 24.8 | 32.3 |
| Early pods | 65.0 | 63.9 | 46.2 | 59.1 | 72.5 | 57.3 | 51.0 | 59.3 |
| Late pods | 2.5 | 1.0 | 0.9 | 1.9 | 2.6 | 3.2 | 2.8 | 2.1 |
| Pods (Total) | 67.5 | 64.9 | 47.1 | 61.0 | 75.1 | 60.5 | 53.8 | 61.4 |
| Flowers and pods (Combined) | 78.7 | 72.6 | 62.1 | 78.0 | 82.8 | 74.4 | 65.8 | 73.5 |
| Seeds | 10.2 | 2.1 | 1.5 | 3.1 | 5.6 | 6.6 | 2.6 | 4.5 |

[a] For description of developmental stages, see Materials and Methods.

2. Patterns of abortion. The patterns of abortion in the G. soja plants examined here were very similar to patterns observed for G. max cultivars (Brevedan et al., 1978; Wiebold et al., 1981; Carlson and Lersten, 1987). The most active abscission periods among the reproductive characters for G. soja were for late flowers (five days post-anthesis) and early pods (less than 2 cm in length) (Table 28). Approximately 85% of all flower abortion events occurred at the late flowering stage; 96% of total pod abortion occurred at the early pod stage (Table 29).

Ninety-five percent (95%) of seed abortion in G. soja occurred at the basal ovule position (42 out of 44 total seed abortion events). The basal ovule is the last one within the pod to be fertilized and may abort more frequently because it must compete with already developing ovules for limited moisture and nutrients from the stem of the inflorescence (Carlson and Lersten, 1987).

Table 29. Proportion of flower and pod abortion events during early and late reproductive developmental stages in the wild soybean[a].

| Population | Flower | | Pod | |
|---|---|---|---|---|
| | Early | Late | Early | Late |
| 1 | 0.291 | 0.709 | 0.963 | 0.037 |
| 2 | 0.200 | 0.800 | 0.985 | 0.015 |
| 3 | 0.091 | 0.909 | 0.980 | 0.020 |
| 4 | 0.038 | 0.962 | 0.969 | 0.031 |
| 5 | 0.086 | 0.914 | 0.965 | 0.035 |
| 6 | 0.153 | 0.847 | 0.947 | 0.053 |
| 7 | 0.213 | 0.787 | 0.948 | 0.052 |
| Overall mean | 0.153 | 0.847 | 0.965 | 0.035 |

[a] Descriptions of reproductive developmental stages are given in Chapter V, Materials and Methods.

## DISCUSSION

The six yield component characters examined for Glycine soja can be separated into two groups containing three characters each: 1) BPI, FPI and IPI (i.e., traits early in the reproductive developmental pathway) and 2) MPI, SPI and SWI (i.e., traits late in the reproductive developmental pathway). Within each of these groups, the characters were significantly correlated; however, no significant correlations were detected between the two groups. The two sets of characters differ with respect to their capacity to discriminate between populations.

### Yield component variation among populations

The results of the ANOVA demonstrated that variation in the yield characters BPI, FPI and IPI was explained, in large part, by the independent variable, population. No significant differences among populations were detected for MPI, SPI and SWI. The differences between the two sets of traits in their capacity to discriminate between the populations may be due to several inter-related factors. First, as the wild soybeans were grown in the greenhouse, they were essentially free from many of the stresses (e.g., herbivory, competition, temperature and moisture fluctuations) and resource limitations often found in natural habitats. Second, the

indeterminate growth habit of the wild soybean and plasticity of the yield characters allows the plant to adjust reproductive output to meet resource availability (Schaal, 1980; Stephenson, 1984). Thus, genetic differences among populations in some characters may have been obscured by a large and positive environmental component (Primack and Antonovics, 1981). A third possibility is that the late yield components are under such strong selective contraints that there is little variation in the genes which condition their development. Fourth, the apparent stability of the late yield component characters across the seven populations may have been mediated through plasticity in physiological characters, which may have been significantly different among populations (Schlichting, 1986). Finally, it is possible that the number of plants per population or number of inflorescences per plant examined were not large enough sample sizes to detect differences among populations for the late yield components.

These results illustrate some of the difficulties encountered when we test for differences among populations for highly plastic morphological traits using small sample sizes and only a single environmental regime.

**Genetic variation and variation in yield components**

In Chapter II, I revealed that populations 1, 2, 5 and 6 (Hatsunedai, Kanodanchi, Asahigaoka-1 and Asahigaoka-2) had higher numbers of alleles per locus, polymorphism and

observed mean heterozygosity compared to populations 3, 4 and 7 (Nishi-asahigaoka, Yata and Kakitagawa). The results here demonstrated that populations 1, 2, 5 and 6 also produced more flower buds, flowers and immature pods per inflorescence than did plants from populations 3, 4 and 7. This trend did not hold as well for MPI, SPI and SWI (perhaps for the reasons outlined above).

In light of these results, the question arises, "What is the relationship between isozymes and the reproductive yield characters?" In an analysis of isozyme and morphological variation in the wild barley (Hordeum spontaneum), Nevo et al. (1979) found no correlation between polymorphism or expected mean heterozygosity and four reproductive characters ( grain weight, and the lengths and widths of grain spikelets, individual grains and the longest glume). These researchers suggested that the lack of correlation indicated that isozymes and spikelet morphology were changing along different evolutionary lines (Nevo et al., 1979).

In the present study, some yield components were significantly correlated with isozyme variation while others were not. No obvious patterns of significant correlations emerged. Additional studies using more populations and larger sample sizes per population are needed before an understanding of the relationship between isozymes and specific morphological traits can be realized.

**Variation in yield components and habitat**

No significant correlations were detected among environmental variables and the population means for the yield component characters. However, without additional physical and historical information about the collection sites, the wild soybean populations, and the degree of phenotypic plasticity in the characters studied, the biological significance of this result for wild soybean population structure in general remains unclear.

**Insights into the reproductive ecology of the soybean**

Rates of abortion of reproductive units observed for G. soja in the present study are in the upper range of rates observed for G. max cultivars (Van Schaik and Probst, 1985b; Hansen and Shibles, 1978; Dominguez and Hume, 1978; Wiebold et al., 1981). The intense artificial selection pressures imposed on the cultigen for high seed weight and number appear to have increased its overall reproductive efficiency by lowering the rates of reproductive abscission. The patterns of abortion, however, remain largely unchanged between the two soybean species. If patterns of reproductive abortion are associated primarily with resource availability, they may have a larger environmental versus genetic component and, therefore, would not be expected to change significantly under selection pressures. Bear in mind that the numbers and patterns described here for G. max are based on data from

both indeterminate and determinate plants, and from plants across different maturity groups. No studies are available for indeterminate, late maturing G. max cultivars, precluding more meaningful comparisons of abortion rates and patterns with the G. soja plants from Japan examined in this study.

Studies on yield in soybean using carbon dioxide and nitrogen enrichment and canopy manipulation suggest that nutrient resources, particularly photosynthates, are the major factors limiting seed set (Harmond and Brun, 1971; Weil and Ohlrogge, 1976; Brevedan et al., 1978; McBlain and Hume, 1981). It seems energetically costly, therefore, for a soybean plant to produce flowers and immature fruits that subsequently abort, effectively removing resources that could have been used by the plant to improve seed quality and/or number. One method used to assess 'cost' of reproductive abscission to overall plant fitness is to quantify the amount of the plant's resources (i.e., biomass, percentage of organic and inorganic nutrients) which are invested in the aborted units. From such studies, researchers generally have concluded that the cost of aborting an early-stage fruit is much smaller than aborting a late-stage fruit (Stephenson, 1981; Bookman, 1983).

The patterns I observed for G. soja are consistent with these analyses, and I propose the following hypothesis to explain the results observed in the present study. Wild soybean plants produce many flower buds because the cost to

the plant (in terms of energy input) if abscission occurs is very small. The abortion rates among flower buds are low because the plant has adequate nutrient resources early in reproductive development to maintain them. At the beginning of the flowering period, resources still are abundant. But as fertilization occurs, and pods and seeds develop, nutrients are depleted. The plants adjust resource allocation to favor fruits and seeds most likely to mature. Usually, these will be undamaged fruits and seeds further along in their development (Lee, 1988). This adjustment of resource allocation when nutrients are critical may explain why I observed such high rates of abortion for late-stage flowers and early-stage pods in the present study. The low rate of seed abortion observed for G. soja makes sense under this hypothesis for two reasons. First, the plant has "fine-tuned" its resource allocation pattern to avoid further investment of nutrients in reproductive units less likely to mature. Second, seed abortion is very costly to a plant in terms of energy input and fitness.

The model of soybean reproductive biology described above does not necessarily apply to all plant species or to all situations (Stephenson, 1984). Also, it deals primarily with 'proximate' causes of abortion (Stephenson, 1981). From an evolutionary standpoint, then, why do flowering plants commonly produce far more reproductive structures than can be maintained to maturity? Several theories have been proposed

to explain this phenomenon (Stephenson, 1981, 1984, 1986; Sutherland, 1987). First, a large number of flowers may serve to attract pollinators to ensure successful pollinations. Second, pollen availability and/or quality may be the limiting factor in fruit-set. Third, excess flowers may increase male fitness through increased pollen donation. Fourth, excess reproductive structures may provide a buffer against herbivory or damage due to abiotic factors. Fifth, by producing more flowers and fruits than can be matured, a plant may regulate offspring quality. Finally, excess reproductive units may allow for a plant to take advantage of years when resources are abundant or to compensate for pollinator density and/or rates of pollinator visitation.

These theories apply mainly to outcrossing plant species. For inbreeders, pollen availability usually is not limiting and pollinators are not essential to successful fertilization (Abernethy et al., 1977; Sutherland, 1987). Only the last three hypotheses (i.e., buffering against reproductive loss, selective abortion and "bet-hedging" in response to resource availability) are likely 'ultimate' explanations for the regular production of excess reproductive structures in primarily inbreeding populations of plants.

# CHAPTER VI

## PROJECT SUMMARY AND APPLICATION TO SOYBEAN GERMPLASM CONSERVATION

### A. Project summary

#### Isozyme and quantitative trait variation

Plants within each of the seven natural populations of G. soja were highly uniform isoenzymatically, but were highly variable phenotypically. The high degree of within-population isozyme uniformity is consistent with the predominantly selfing mating system of this species, the severely restricted gene flow among populations and the apparent small effective population sizes. The variability in the soybean phenotype may have been due, in part, to phenotypic plasticity, although additional studies are needed to verify this.

The wild soybean populations from which seed were collected in the field were small in size and isolated (H.I. Oka, personal communication);therefore, the rates of gene flow among the soybean populations were likely very low. Thus, the patterns of population differentiation observed for isozymes and quantitative traits may have occurred according to Sewall Wright's "island model" (Wright, 1969). According to Wright's model, the extent of differentiation of gene frequencies among populations is high when the

product of effective population size ($N_e$) and migration (m) is small. Under such conditions, selection and genetic drift occur more or less independently in each population. Thus, the electrophoretic and genetically-based (as opposed to environmentally-based) phenotypic differences among the soybean populations are probably due to a combination of genetic adaptations to local environmental conditions and stochastic processes, such as genetic drift.

Although the amount of information available on environmental variation at the soybean seed collection sites was quite limited, one component, the percentage of the collection site occupied by G. soja relative to other co-habitating plant species (COVER; see Appendices I and II), was significantly correlated with genetic diversity (A, P95, P99, $H_{obs}$) and phenotypic diversity ($D^2$) estimates. Thus, inter- and intra-specific plant associations may have played a role in shaping the genetic structure of the seven populations of G. soja. Plant associations have been indicated as a significant contributing factor to population structure in the annual grass, Anthoxanthum (Kiang, 1982; Antonovics et al., 1988) and deserve consideration in future studies of G. soja.

The degree of concordance between isozymes and quantitative traits in describing population differentiation was statistically significant in 1987, but not in 1986. The inconsistency probably resulted from a combination of

heterogeneity in the data sets due to small sample sizes, experimental error and the confounding effects of phenotypic plasticity.

## Reproductive biology of the wild soybean

Among the six reproductive yield components examined for the wild soybean, the three early components (number of flower buds, flowers and immature pods per inflorescence) were significantly different among populations. The three late-yield components (number of mature pods and mature seeds per inflorescence, and mature seed weight per inflorescence) were not significantly different among the populations. The variability detected among early-yield components may provide the wild soybean flexibility in reproductive effort in the face of varying resource availability. The lack of significant variation among populations in the late yield components may reflect the resource dependent nature of these traits or the lack of variation in the genes which condition the development of these traits.

The rates of abortion of flower buds, flowers and pods observed for G. soja were in the upper ranges of rates observed for G. max cultivars (Wiebold et al., 1981). The patterns of abortion in G. soja were very similar to those of the cultigen. For the wild soybean, 85% of the flower abortion activity occurred one to seven days post-anthesis

and the 95% of pod abortion occurred when pods were immature and less than 2 cm long. Approximately 96% of the seed abortion events occurred at the basal ovule, the position within the pod which is the last to be fertilized. From these results, I concluded that 1) artificial selection for high seed weight and yield has increased the overall reproductive efficiency in G. max by lowering the rate of reproductive abscission in this species, demonstrating the existence of a genetic component to reproductive efficiency and 2) the patterns of reproductive abscission in the wild soybean likely are associated with resource availability and probably do not have a significant genetic component.

## Toward future research

The results of this research highlighted several areas in which further study is needed. I believe that two of these areas, gene flow and phenotypic plasticity, are particulary critical to our understanding of the population biology of G. soja. Gene flow is one of the major forces shaping genetic structure in all organisms, yet it is one of the most elusive processes to measure (Slatkin, 1987). Two important questions are, 1) how temporally and spatially variable is the effective population size in G. soja and 2) what factors (e.g. population shape and density, community structure) affect the rates of outcrossing? Fortunately, recent advances in molecular biology have provided us with

tools which will be invaluable in demographic studies of plant populations. Regions of highly variable "minisatellite" nucleotide sequences are dispersed throughout the genome of many organisms and have been used successfully as DNA "fingerprints" of individuals in studies of parentage in house sparrow populations (Wetton et al., 1987; Burke and Bruford, 1987). The application of this technology to plant populations will enhance our capacity to quantify gene flow in the field.

The second critical area for further research is the nature of phenotypic plasticity. Presumably, phenotypic plasticity is an important "buffering" mechanism for individual survival in the face of environmental variation when a population has a rather uniform gene pool (Moran and Marshall, 1978; Wu and Jain, 1978). We need to quantify the amount of phenotypic plasticity in natural populations of wild soybean to better understand its role in conditioning the capacity of this species to adapt to so many diverse environmental conditions. As discussed in Chapter III, phenotypic plasticity can be assessed by measuring traits across several environments and/or using transplant studies (Schlichting, 1986).

## B. Germplasm conservation in the soybean

To collect, manage and maintain genetically diverse germplasm resources efficiently, we must first understand

the genetic structure of populations; i.e., we must know how much genetic diversity exists and how it is apportioned within and between populations. If the objective of a germplasm conservation program is to collect and preserve as much genetic diversity as possible, Marshall and Brown (1975) have argued that alleles with a high frequency (10% and greater) and with restricted or localized distributions merit high priority in a collection plan. They also recommended that between 50 and 100 individuals be collected from as many ecologically different sites within a target area as possible given the limitations of time, money, personnel and politics. This sampling strategy assumes that at least some of the observed localized differentiation of alleles is adaptive.

In the present study, the populations were clustered into three or four groups by both canonical discriminant analyses (Chapters II and III) and cluster analyses (Chapter IV). However, group membership in the clusters varied depending on the analysis, the character type (isozymes versus phenotypic traits) and, for the phenotypic data, the year. Thus, the adaptive nature of the isozyme and phenotype differences among the soybean populations, and the reliability of the population clusters are questionable. While multivariate analyses are useful for germplasm conservation efforts, they clearly should not be the sole criteria on which collection decisions are made (See

Oldfield, 1989).

Given the variability of the wild soybean phenotype, isozymes are likely to be the most reliable indicators of differences among populations in the present study. Therefore, collections from the population clusters based on either canonical discriminant analysis or cluster analysis of isozyme variation would be sufficient to represent most of the genetic diversity present among the seven populations. Although the clusters generated from CDA and cluster analysis were somewhat different, most of the differences among the soybean populations in the present study were due to differences in allele frequencies, not in the presence or absence of alleles. Thus, most of the alleles would be represented in a sample regardless of which grouping method was used as an aid to collecting.

If individual buffering (phenotypic plasticity) plays a larger role in adaptation than population buffering (differential fitness of various co-existing genotypes), then the sampling scheme described by Marshall and Brown (1975) may be inappropriate for large-scale collections of G. soja. Instead, collections should be made randomly or based on other criterion, such as population size or evidence of agronomically useful traits (Brown, 1978). Further studies of natural populations of wild soybean are needed before a comprehensive collection strategy can be developed.

## Biotechnology in soybean germplasm conservation

The most common approach to using germplasm resources is to incorporate genes into the genome of a cultivar (or potential cultivar) through repeated backcrosses (Fehr, 1987). For soybeans this process is tedious and time consuming because the weedy characteristics of G. soja are not the characteristics of a successful soybean cultivar (Fehr, 1987). Advances in DNA transfer technology have provided promising evidence for the use of genetic engineering for soybean improvement. Stable transformants of G. max cultivars recently were obtained using particle acceleration of DNA-coated gold pellets into meristems of immature soybean seedlings (McCabe et al., 1988). Theoretically, a gene from any species could be transferred into the soybean genome using this technology.

New technologies must not be viewed as an alternative to germplasm collection and preservation, but as a part of a multifaceted approach to create and maintain genetic reserves. Indeed, the success of biotechnology in plant breeding depends on the availability of novel genes and gene combinations and may not be as effective as traditional breeding methods for improvement of quantitative traits, such as yield (Oldfield, 1989).

## Beyond practical applications of gene conservation

The potential agronomic usefulness of the wild soybean

populations examined in the present study was not assessed. If they do not contain genetic variation which is useful for current soybean breeding programs and goals, why should we seek to characterize and maintain the genetic diversity within these populations? One reason is that natural populations have great instructive value for studying evolutionary processes, such as the formation and maintainence of co-adapted gene complexes and the development of community structure (Frankel, 1974). A second reason is that the long-term value of genetic resources cannot be determined based on present-day needs and concerns. As the human population grows and arable land diminishes, our agronomic needs and practices will change and we will look to genetic reserves for the raw material to meet these challenges. Finally, beyond the practical economic and social aspects of germplasm conservation, is the concept of our "evolutionary responsibility" (Frankel, 1974). Like air, water and soil, genes are precious natural resources and are part of our global "evolutionary heritage" (Frankel, 1974). We need to study, protect and conserve genetic resources to ensure ecological diversity and balance for future generations of all species.

## LITERATURE CITED


Abernethy, R.H, R.G. Palmer, R.M. Shibles and I.C. Anderson. 1977. Histological observations on abscising and retaining soybean flowers. Can. J. Plant Sci. 57:713-716.

Adams, M.W. 1967. Basis of yield component compensation in crop plants with special reference to the field bean, Phaseolus vulgaris. Crop Sci. 7:505-510.

Allard, R.W., S.K. Jain and P.L Workman. 1968. The genetics of inbreeding populations. Adv. Genet. 14:55-131.

Allard, R.W. 1988. Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. J. Hered. 79:225-238.

Antonovics, J. and A.D. Bradshaw. 1970. Evolution in closely adjacent plant populations. VIII. Clinal patterns at a mine boundry. Heredity 25:349-362.

Antonovics, J., A.D. Bradshaw and R.G. Turner. 1971. Heavy metal tolerance in plants. Adv. Ecol. Res. 7:1-85.

Antonovics, J., N.C. Ellstrand and R.N. Brandon. 1988. Genetic variation and environmental variation: expectations and experiments. IN: (L.D. Gottlieb and S.K. Jain, eds.) Plant Evolutionary Biology. Chapman and Hall. pp. 275-303.

Atchley, W.R. 1971. A statistical analysis of geographic variation in the pupae of three species of Culicoides (Diptera: Ceratopogonidae). Evolution 25:51-74.

Aung, L.H., G.R. Buss and J.M. Byrne. 1986. Developmental characteristics of two differential yielding soybean genotypes. Phyton 46:77-86.

Ayala, F.J. 1982a. The genetic structure of species. IN: Perspectives on evolution (R. Milkman, ed.). Sinauer Assoc., Inc. (Sunderland, MA). pp. 60-82.

Ayala, F.J. 1982b. Genetic variation in natural populations: Problem of electrophoretically cryptic alleles. Proc. Nat. Acad. Sci., USA 79:550-554.

Beremand, M.H. 1975. Soybean ADH: Dimeric structure and subunit composition of multiple isozymes. Abst. Int. Cong. Genet. Genetics 80:512.

Bookman, S.S. 1983. Costs and benefits of flower abscission and fruit abortion in Asclepias speciosa. Ecology 64: 264-273.

Bradshaw, A.D. 1965. Evolutionary significance of phenotypic plasticity in plants. Adv. Genet. 13:115-155.

Bradshaw, A.D. 1972. Some of the evolutionary consequences of being a plant. Evolutionary Biology 5:25-47.

Brevedan, R.E., D.B. Egli and J.E. Leggett. 1978. Influence of N nutrition on flower and pod abortion and yield of soybeans. Agron. J. 70:81-84.

Brewbaker, J.L., M.D. Upadhya, Y. Makinen and T. MacDonald. 1968. Isozyme polymorphisms in flowering plants. III. Gel electrophoretic methods and applications. Physiol. Planta. 21:930-940.

Brown, A.H.D., D.R. Marshall and L. Albrecht. 1974. The maintenance of alcohol dehydrogenase polymorphism in Bromus mollis. L. Aust. J. Bot. Sci. 27:545-559. and Forest Insects.

Brown, A.H.D. 1978. Isozymes, plant population genetic structure, and genetic conservation. Theor. Appl. Genet. 52: 145-157.

Brown, A.H.D. 1979. Enzyme polymorphism in plant populations. Theor. Pop. Biol. 15:1-42.

Brown, A.H.D. and B.S. Weir. 1983. Measuring genetic variability in plant populations. IN: Isozymes in plant genetics and breeding, Part A. (S.D. Tanksley and T.J. Orton, eds.) Elsevier Science Publishers (New York) pp. 219-239.

Bryant, E.H. 1984. A comparison of electrophoretic and morphometric variability in the face fly, Musca autumnalis. Evolution 38:455-459.

Bult, C.J. and Y.T. Kiang. 1989. Inheritance and genetic linkage tests of an esterase locus in the cultivated soybean, Glycine max. J. Hered. 80:82-85.

Bult, C.J., Y.T. Kiang, Y.C. Chiang, H.J.Y. Doong and M.B. Gorman. 1989. Electrophoretic methods for soybean genetics studies. Soybean Genetics Newsletter 16:175-187.

Burke, T. and M.W. Bruford. 1987. DNA fingerprinting in birds. Nature 327:149-152.

Buttery, B.R. and R.I. Buzzell. 1968. Peroxidase activity in seeds of soybean varieties. Crop Sci. 8:722-725.

Buttery, B.R. and R.I. Buzzell. 1971. Properties and inheritance of urease isoenzymes in soybean seeds. Can. J. Bot. 49:1101-1105.

Buzzell, R.I. and B.R. Buttery. 1969. Inheritance of peroxidase activity in soybean seed coats. Crop Sci. 9: 387-388.

Cardy, B.J., C.W. Stuber and M.M. Goodman. 1980. Techniques for starch gel electrophoresis of enzymes from maize (Zea mays L.). Mimeograph series #1317. Dept. of Statistics. N.C.S.U. Raleigh, North Carolina 27650. 31 pp.

Carlson, J.B. and N.R. Lersten. 1987. Reproductive morphology. IN: Soybeans: Improvement, production and uses. 2nd ed. (J.R. Wilcox, ed.) Agronomy Monograph No. 16. pp. 95-134.

Carpenter, J.A. and W.A. Fehr. 1986. Genetic variability for desirable agronomic traits in populations containing Glycine soja. Crop Sci. 26:681-686.

Chiang, Y.C. 1985. Genetic and quantitative variation in wild soybean (Glycine soja) populations. Ph.D. Dissertation, University of New Hampshire.

Chiang, Y.C. and Y.T. Kiang. 1987. Inheritance and linkage relationships of 6-phosphogluconate dehydrogenase isozymes in soybean. Genome 29:786-792.

Chiang, Y.C., M.B. Gorman and Y.T. Kiang. 1987. Inheritance and linkage analysis of phosphoglucose isomerase isozymes in soybeans. Biochem. Genet. 25:893-900.

Chiang, Y.C. and Y.T. Kiang. 1988. Genetic analysis of mannose 6-phosphate isomerase in soybeans. Genome 30:808-811.

Clausen, J., D.D. Keck and W.M. Hiesey. 1940. Experimental studies on the nature of species. I. Effect of varied environments on wester North American plants. Carnegie Inst. Wash. Publ. 520:1-452.

Clausen, J., D.D. Keck and W.M. Hiesey. 1948. Experimental studies on the nature of species. III. Environmental response of climatic races of Achillea. Carnegie Inst. Wash. Publ. 581:1-129.

Clay, K. and D.A. Levin. 1989. Quantitative variation in Phlox: comparison of selfing and outcrossing species. Amer. J. Bot. 76:577-588.

Clegg, M.T. and R.W. Allard. 1972. Patterns of genetic differentiation in the slender wild oat species, Avena barbata. Proc. Natl. Acad. Sci. USA 69:1820-1824.

Conkle, M.T., P.D. Hodgkiss, L.B. Nunnaly and S.C. Hunter. 1982. Starch gel electrophoresis of conifer seeds: A laboratory manual. USDA General Technical Report PSW-64.

Cook, R.E. 1975. The photoinductive control of seed weight in Chenopodium rubrum L. Amer. J. Bot. 62:427-431.

Darmency, H. and J. Gasquez. 1983. Esterase polymorphisms and genetic growth form differentiation in Poa annua L. New. Phytol. 95:289-297.

Delannay, X., D.M. Rogers and R.G. Palmer. 1983. Relative genetic contributions among ancestral lines to North American soybean cultivars. Crop Sci. 23:944-949.

Dixon,M. and E.C. Webb. 1979. Enzymes, 3rd ed. Academic Press, Inc. (New York).

Dobzhansky, T. 1970. Genetics of the evolutionary process. Columbia University Press (New York).

Domingo, W.E. 1945. Inheritance of number of seeds per pod and leaflet shape in the soybean. J. Agric. Res. 70: 251-268.

Dominguez, C. and Hume, D.J. 1978. Flowering, abortion and yield of early-maturing soybeans. Agron. J. 70:801-805.

Doong, J.Y.H. 1986. An electrophoretic study on soybean isozymes. M.S. Thesis, University of New Hampshire.

Doong, J.Y.H. and Y.T. Kiang. 1987a. Inheritance of aconitase isozymes in soybean. Genome 29:713-717.

Doong, J.Y.H. and Y.T. Kiang. 1987b. Inheritance of soybean endopeptidase. Biochem. Genet. 25:847-853.

Doong, J.Y.H. and Y.T. Kiang. 1988. Inheritance study on a soybean fluorescent esterase. J. Hered. 79:399-400.

Doyle, J.J. and R.N. Beachy. 1985. Ribosomal gene variation in soybean (Glycine) and its relatives. Theor. Appl. Genet. 70:369-376.

Doyle, J.J. 1988. 5S ribosomal gene variation in the soybean and its progenitor. Theor. Appl. Genet. 75:621-624.

Duvick, D.N. 1984. Genetic diversity in major farm crops on the farms and in reserve. Econ. Bot. 38:161-178.

Erickson, L.R., H.D. Boldeng and W.D. Beversdorf. 1981. Early generation selection for protein in Glycine max X Glycine soja crosses. Can. J. Plant Sci. 61:901-908.

Ertl, D.S. and W.R. Fehr. 1985. Agronomic performance of soybean genotypes from Glycine max X Glycine soja crosses. Crop Sci. 25:589-592.

Fairlet, J.L. and G.L. Kilgour. 1966. Essentials of biological chemistry. Reinhold Pub. Co.

Farris, J.S. 1973. On comparing the shapes of taxonomic trees. Syst. Zool. 22:50-54.

Fehr, W.G., C.E. Caviness, D.T. Burmood and J.S. Pennington. 1971. Stage of development descriptions for soybeans, Glycine max (L.) Merrill. Crop Sci. 11:929-931.

Fehr, W.G. 1987.Soybean. IN: Principles of cultivar development. Vol.2: Crop species. (W.G.Fehr, ed.) MacMillan Pub. Co. (New York) pp. 533-576.

Ferrer-Monge, J.A. 1974. Esterase isozyme patterns in Glycine max exposed to gamma radiation. Can. J. Bot. 52:273-275.

Frankel, O. 1974. Genetic conservation:our evolutionary responsibility. Genetics 78:53-65.

Frankel, O.H. and M.E. Soule. 1981. Conservation and Evolution. Cambridge University Press (Cambridge). pp. 232-244.

Giles, B.E. and K.J.R. Edwards. 1983. Quantitative variation within and between populations of wild barley, Hordeum murinum. Heredity 51:325-333.

Giles, B. 1984. A comparison between quantitative and biochemical variation in the wild barley, Hordeum murinum. Evolution 38:34-42.

Goodwin, T.W. and E.I. Mercer. 1983. Introduction to Plant Biochemistry, 2nd ed. Pergamon Press (New York).

Gorman, M.B. and Y.T. Kiang. 1977. Variety-specific electrophoretic variants of four soybean enzymes. Crop Sci. 17:963-965.

Gorman, M.B. and Y.T. Kiang. 1978. Models for the inheritance of several variant soybean electrophoretic zymograms. J. Hered. 69:255-258.

Gorman, M.B. 1983. An electrophoretic analysis of the genetic variation in the wild and cultivated soybean germplasm. Ph.D. Dissertation, University of New Hampshire.

Gorman, M.B., Y.T. Kiang, R.G. Palmer and Y.C. Chiang. 1983. Inheritance of soybean electrophoretic variants. Soybean Genetics Newsletter 10:67-84.

Gottlieb, L.D. 1971. Gel electrophoresis: A new approach to the study of evolution. BioScience 21:939-944.

Gottlieb, L.D. 1981. Electrophoretic evidence and plant populations. Progr. Phytochem. 7:1-46.

Graef, G.L., W.R. Fehr and C.R. Cianzio. 1989. Relation of isozyme genotypes to quantitative characters in soybean. Crop Sci. 29:683-688.

Greenlee, J.K. and K.S. Rai. 1986. Population differentiation in Collinsia verna (Scrophulariaceae): A multifaceted approach. Genetica 71:51-61.

Griffin, J.D. and R.G. Palmer. 1987. Inheritance and linkage studies with five isozyme loci in soybean. Crop Sci. 27:885-892.

Hadley, H.H. and T. Hymowitz. 1987. Speciation and cytogenetics. IN: Soybeans: Improvement, production and uses, 2nd ed.(J.R. Wilcox, ed.). Agronomy Monograph No. 16. American Society of Agronomy, Inc. (Madison, Wisc.)

Hamrick, J.L. and R.W. Allard. 1975. Correlation between quantitative characters and enzyme genotypes in Avena barbata. Evolution 29:438-443.

Hamrick, J.L., Y.B. Linhart and J.B. Mitton. 1979. Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Ann. Rev. Ecol. Syst. 10:173-200.

Hamrick, J.L. 1982. Plant population genetics and evolution. Amer. J. Bot. 69:1685-1693.

Hansen, W.R. and R. Shibles. 1978. Seasonal log of the flowering and podding activity of field-grown soybeans. Agron. J. 70:47-50.

Harlan, J.R. 1976. Genetic resources of wild relatives of crops. Crop Sci. 16:329-333.

Harmond, L.L. and W.A. Brun. 1971. Effects of atmosphere carbon dioxide enrichment at different developmental stages on growth and yield components of soybeans. Crop Sci. 11:886-888.

Harper, J.L., P.H. Lowell and K.G. Moore. 1970. The shapes and sizes of seeds. Ann. Rev. Ecol. Syst. 1:327-356.

Hartl, D.L. 1980. Principles of population genetics. Sinauer Assoc., Inc. (Sunderland, MA.).

Hedrick, P.W. 1986. Genetic polymorphism in heterogeneous environments: A decade later. Ann. Rev. Ecol. Syst. 17:535-566.

Heslop-Harrison, J. 1964. Forty years of genecology. Adv. Ecol. Res. 2:159-247.

Hildebrand, D.R., J.H. Orf and T. Hymowitz. 1980. Inheritance of an acid phosphatase and its linkage with the Kunitz trypsin inhibitor in seed protein of soybeans. Crop Sci. 20:83-85.

Hubby, J.L. and R.C. Lewontin. 1966. A molecular approach to the study of genic heterozygosity in natural populations. I.The number of alleles at different loci in Drosophila pseudoobscura. Genetics 54:577-594.

Hymowitz, T.A. 1970. On the domestication of the soybean. Econ. Bot. 24:408-421.

Hymowitz, T. and H.H. Hadley. 1972. Inheritance of a trypsin inhibitor variant in seed protein of soybeans. Crop Sci. 12:197-198.

Hymowitz, T. and N. Kaizuma. 1981. Soybean seed protein electrophoresis profiles from 15 Asian countries or regions: Hypotheses of paths of dissemination of soybeans from China. Econ. Bot. 35:10-23.

Jain, S.K. and A.D. Bradshaw. 1966. Evolutionary divergence among adjacent plant populations. I. The evidence and its theoretical analysis. Heredity 21:407-441.

Jain, S.K. 1978. Adaptive strategies: polymorphism, plasticity and homeostasis. IN: (O.T. Solbrig, S.K. Jain, G.B. Johnson and P.H. Raven, eds.). Plant Population Biology. Cornell Univ. Press pp. 160-187.

Kahler, A.L., R.W. Allard, M. Krzakowa, C.F. Wehrhahn and E. Nevo. 1980. Associations between isozyme phenotypes and environment in the slender wild oat (Avena barbata) in Israel. Theor. Appl. Genet. 56:31-47.

Kaizuma, N. and J. Fukui. 1974. Intraspecific differences in contents of seed protein and sulphur-containing amino acids in soya bean (Glycine soja Sieb. & Zucc.) and their significance for plant breeding. Jap. J. Breeding (Isushugaku Zasshi) 24:65-72.

Kaizuma, N., Y. Kiuch and F. Ono. 1980. Breeding of high protein soybeans from a species cross between soybeans Glycine max and wild soybeans Glycine soja: Necessity of repeated backcrossing method. J. Fac. Agric. Iwate Univ. 15:11-28.

Kerr, P.S., D.G. Blevins, B.J. Rapp, D.D. Randall. 1983. Soybean leaf urease: comparison with seed urease. Physiol. Plant. 57:239-345.

Kiang, Y.T. 1981. Inheritance and variation of amylase in cultivated and wild soybeans and their wild relatives. J. Hered. 72:382-386.

Kiang, Y.T. 1982. Local differentiation of Anthoxanthum odoratum L. populations on roadsides. Am. Mid. Nat. 107:340-350.

Kiang, Y.T. and M.B. Gorman. 1983. "Soybean". IN: Isozymes in plant genetics and breeding, Part B. (Tanksley,S.D. and T.J. Orton, eds.) Elsevier Science Pub (Amsterdam) pp. 295-328.

Kiang, Y.T. and M.B. Gorman. 1985. Inheritance of NADP-active isocitrate dehydrogenase isozymes in soybeans. J. Hered. 76:279-284.

Kiang, Y.T., M.B. Gorman and Y.C. Chiang. 1985. Genetic and linkage analysis of a leucine aminopeptidase in wild and cultivated soybean. Crop Sci. 25:319-321.

Kiang, Y.T. 1987. Mapping three protein loci on a soybean chromosome. Crop Sci. 27:44-46.

Kiang, Y.T. and Y.C. Chiang. 1987. Genetic analysis of the Lap2 locus in soybeans and tests of linkage with other loci. J. Hered. 78:131-132.

Kiang, Y.T., Y.C. Chiang and C.J. Bult. 1987. Genetic study of glutamate oxaloacetic transaminase in soybean. Genome 29:370-373.

Kiang, Y.T., Y.C. Chiang, J.Y.H. Doong and M.B. Gorman. 1987. Genetic variation of soybean germplasm. IN: Crop exploration and utilization of genetic resources. pp. 93-99.

Kiang, Y.T. and Y.C. Chiang. 1989. Geographical variation in the reproductive biology of wild soybean populations. IN:(J.H. Bock and Y.B. Linhart, eds.) The Evolutionay Ecology of Plants. Westview Press. pp. 469-489.

Kidd, G.H. 1988. Engineered soybean becomes a reality: impact on agribusiness. Biotechnology 6:870.

Kloth, R.H. and T. Hymowitz. 1985. Re-evaluation of the inheritance of urease in soybean seed. Crop Sci. 25: 352-354.

Kloth, R.H., J.C. Polacco and T. Hymowitz. 1987. The inheritance of a urease-null trait in soybeans. Theor. Appl. Genet. 73:410-418.

Koehn, R.K., R.I. Newell and F. Immermann. 1980. Maintenance of an aminopeptidase allele frequency cline by natural selection. Proc. Natl. Acad. Sci. USA 77:5385-5389.

Lande, R. 1977. The influence of the mating systems on the maintenance of genetic variability in polygenic characters. Genetics 86:485-498.

Lande, R. 1982. A quantitative genetic theory of life-history evolution. Ecology 63:607-615.

Lande, R. and S.J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37: 1210-1226.

Lawrence, P.K. and K.J. Frey. 1975. Backcross variability for grain yield in oat species crosses (Avena sativa X A. sterilis L.) Euphytica 24:77-85.

Lee, T.D. 1988. Patterns of fruit and seed production. IN: Reproductive strategies of plants (J. and L. Lovett-Doust, eds.) CRC Press (New York). pp. 179-202.

Lehninger, A.L. 1982. Principles of biochemistry. Worth Publishers, Inc. (New York). 1011 pp.

Levin, D.A. 1975. Genic heterozygosity and protein polymorphism among local populations of Oenothera biennsis. Genetics 79:477-491.

Lewontin, R.C. and J.L. Hubby. 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of Drosophila pseudoobscura. Genetics 54:595-609.

Lewontin, R.C. 1966. On the measurement of relative variability. Syst. Zool. 15:141-147.

Lewontin, R.C. 1974. The genetic basis of evolutionary change. Columbia University Press (New York). 345 pp.

Lewontin, R.C. 1984. Detecting population differences in quantitative characters as opposed to gene frequencies. Am. Nat. 123:115-124.

Linhart, Y.B. 1974. Intra-population differentiation in annual plants. I. Veronica perigrina L. raised under non-competitive conditions. Evolution 28:232-243.

Linhart, Y.B., M.C. Davis and J.B. Mitton. 1981. Genetic control of allozymes of shikimate dehydrogenase in Ponderosa pine. Biochem Genet. 19:641-646.

Little, T.M. and F.J. Hills. 1978. Agricultural experimentation: Design and analysis. J. Wiley and Sons (New York): 350 pp.

Lotz, L.A.P. and D.H. Spoorsmakers. 1988. Differentiation in reproductive characteristics within a populations of Plantago major L. ssp. pleiosperma Pilger. Acta Oecologia (Oecologia Plantarium) 9:11-18.

Loveless, M.D. and J.L. Hamrick. 1984. Ecological determinants of genetic structure in plant populations. Ann. Rev. Ecol. Syst. 15:65-95.

Luedders, V.D. 1977. Genetic improvement in yield of soybeans. Crop Sci. 17:971-977.

Manchenko, G.P. 1988. Subunit structure of enzymes: Allozymic data. Isozyme Bull. 21:144-158.

Marshall, D.R. and R.W. Allard. 1970. Isozyme polymorphism in natural populations of Avena fatua and A. barbata. Heredity 25:373-382.

Marshall, D.R. and A.H.D. Brown. 1975. Optimum sampling strategies in genetic conservation. IN: Crop genetics resources for today and tomorrow (O.H. Frankel and J.G. Hawkes, eds.) Cambridge Univ. Press. pp. 53-80.

Marshall, D.R. , N.L. Fowler and D.A. Levin. 1985. Plasticity in yield components in natural populations of three species of Sesbania. Ecology 66:753-761.

McBlain, B.A. and D.J. Hume. 1981. Reproductive abortion, yield components and nitrogen fixation in three early soybean cultivars. Can. J. Plant Sci. 61:499-505.

McCabe, D.E., W.F. Swain, B.J. Martinell and P. Christou. 1988. Stable transformation of soybean (Glycine max) by particle acceleration. Biotechnology 6:923-927.

McWilliams, E.L., R.Q. Landers and J.P. Mahlstede. 1968. Variation in seed weight and germination in populations of Amaranthus retroflexus L. Ecology 49:290-296.

Moran, G.F. and D.R. Marshall. 1978. Allozyme uniformity within and variation between races of the colonizing species Xanthium strumarium L. (Noogoura Burr). Aust. J. Biol. Sci. 31:283-291.

Morita, Y., F. Yagi, S. Aibara and H. Yamashita. 1976. Chemical composition and properties of soybean beta-amylase. J. Biochem. 79:591-603.

Morris, P.C., R.C. Miller and D.J. Bowles. 1985. Endopeptidase activity in dry harvest-ripe wheat and barley grains. Plant Sci. 39:121-124.

Moss, D.W. 1982. Isoenzymes. Chapman and Hall (New York). 204 pp.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci., USA 70:3321-3323.

Nei, M. 1978. Estimation of average heterozygosity of genetic distance from a small number of individuals. Genetics 89:583-590.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press (New York). 512 pp.

Nevo, E. 1978. Genetic variation in natural populations: Patterns and theory. Theor. Pop. Biol. 13:121-177.

Nevo, E., D. Zohary, A.H.D. Brown and M. Haber. 1979. Genetic diversity and environmental associations of wild barley, Hordeum spontaneum, in Israel. Evolution 33: 815-833.

Nevo, E., E. Golenberg and A. Beiles. 1982. Genetic diversity and environmental associations of wild wheat, Triticum dicoccoides, in Israel. Theor. Appl. Genet. 62:241-254.

Nevo, E. 1987. Plant genetic resources: Prediction by isozyme markers and ecology. IN: Current topics in biological and medical research. Vol. 16: Agriculture, physiology and medicine pp. 247-267.

Norusis, M.J. 1985. SPSS-X: Advanced statistics guide. McGraw-Hill Book Co. (New York). 505 pp.

Oldfield, M.L. 1989. The value of conserving genetic resources. Sinauer Assoc., Inc. 379 pp.

O'Malley, D., N.C. Wheeler and R.P. Guries. 1980. A manual for starch gel electrophoresis. Staff paper series 11. Madison, WI. Dept. of Forestry, College of Agriculture and Life Sciences. University of Wisconsin. 16 pp.

Orf, J.H. and T. Hymowitz. 1979. Inheritance of the absence of the Kunitz trypsin inhibitor in seed protein of soybeans. Crop Sci. 19:107-109.

Palmer, R.G., K.L. Newhouse, R.A. Graybosch and X. Delannay. 1987. Chromosome structure of the wild soybean. J. Hered. 78:243-247.

Parker, M.A. 1988. Disequilibrium between disease-resistance variants and allozyme loci in an annual legume. Evolution 42:239-247.

Plucknett, D.L., N.J.H. Smith, J.T. Williams and N.M. Anishetty. 1987. Gene banks and the world's food. Princeton University Press (Princeton).

Polacco, J.C. and E.A. Havir. 1979. Comparisons of soybean urease isolated from seed and tissue culture. J. Biol. Chem. 254:1707-1715.

Price, S.C., K.M. Shumaker, A.L. Kahler, R.W. Allard and J.E. Hill. 1984. Estimates of differentiation obtained from enzyme polymorphisms and quantitative characters. J. Hered. 75:141-142.

Primack, R.B. 1978. Regulation of seed yield in the genus Plantago. J. Ecol. 66:835-847.

Primack, R.B. and J. Antonovics. 1981. Experimental ecological genetics in Plantago. V. Components of seed yield in the ribwort plantain Plantago lanceolata L. Evolution 35:1069-1080.

Raboy, V., D.B. Dickinson and F.E. Below. 1984. Variation in seed total phosphorous, phytic acid, zinc, calcium, magnesium, and protein among lines of Glycine max and G. soja. Crop Sci. 24:431-434.

Rai, K.N. and S.K. Jain. 1982. Population biology of Avena. IX. Gene flow and neighborhood size in relation to micro-geographic variation in Avena barbata. Oecologia 53:399- 405.

Ram, H.H., Pushpendara, K. Singh and V.D. Verma. 1984. New breeding lines of soybean having a gene for resistance to yellow-mosaic virus from Glycine soja. Indian J. Agric. Sci. 54:1027-1029.

Ramshaw, J.A.M., J.A. Coyne and R.C. Lewontin. 1979. The sensitivity of gel electrophoresis as a detector of genetic variation. Genetics 93:1019-1037.

Ryu, T.B. 1982. Genetic structure of Pinus strobus on foliar isozymes from 27 provenances. Ph.D. Dissertation. Univ. of New Hampshire. Durham, N.H.

SAS, Institute. 1985a. SAS User's Guide:Basics, Version 5 Edition. SAS Institute, Cary, North Carolin 27511.

SAS Institute. 1985b. SAS User's Guide:Statistics, Version 5 Edition. SAS Institute, Cary, North Carolina 27511.

Schaal, B.A. 1980. Reproductive capacity and seed size in Lupus texensis. Amer. J. Bot. 67:703-709.

Schimpf, D.J. 1977. Seed weight of Amaranthus retroflexus in relation to growing season. Ecology 58:450-453.

Schlichting, C. 1986. The evolution of phenotpyic plasticity in plants. Ann. Rev. Ecol. Syst. 17:667-693.

Schoen, D.J. 1982. Genetic variation and the breeding system of Gilia achilleifolia. Evolution 36:361-369.

Schwaegerle, K.E., K. Garbutt and F.A. Bazzaz. 1986. Differentiation among nine populations of Phlox. I. Electrophoretic and quantitative variation. Evolution 40:506-517.

Shaw, C.R. and R. Prasad. 1969. Starch gel electrophoresis of enzymes: A compilation of recipes. Biochem. Genet. 4:297-320.

Shoemaker, R.C., P.M. Hatfield, R.G. Palmer and A.G. Atherly. 1986. Chloroplast DNA variation in the genus Glycine subgenus Soja. J. Hered. 77:26-30.

Silvertown, J. 1989. The paradox of seed size and adaptation. Trends in Ecology and Evolution 4:24-26.

Sisson, V.A., C.A. Brim and C.S. Levings III. 1978. Characterization of cytoplasmic diversity in soybeans by restriction endonuclease analysis. Crop Sci. 18: 991-996.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787-792.

Snaydon, R.W. and M.S. Davies. 1972. Rapid population differentiation in a mosaic environment. II. Morphological variation in Anthoxanthum odoratum. Evolution 26:390-405.

Sokal, R.R. and F.J. Rohlf. 1969. Biometry: The principles and practice of statistics in biological research. W.H. Freeman and Company. 776 pp.

Soltis, P.S. and D.E. Soltis. 1987. Population structure and estimates of gene flow in the homosporous fern, Polysticum munitum. Evolution 41:620-629.

Soybean Research Advisory Institute. 1984. U.S. soybean production and utilization research.

SPSS-X, Inc. 1986. SPSS-X User's Guide, 2nd ed. SPSS-X, Inc. Chicago, Illinois 60611. 988 pp.

Stephenson, A.G. 1981. Flower and fruit abortion: Proximate causes and ultimate functions. Ann. Rev. Ecol. Syst. 12:253-279.

Stephenson, A.G. 1984. The cost of over-initiating fruit. Amer. Mid. Nat.112:379-386.

Stephenson, A.G. 1986. Lotus corniculatus regulates offspring quality through selective fruit abortion. Evolution 40:453-458.

Stuber, C.W., M.M Goodman and R.H. Moll. 1982. Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. Crop Sci. 22:737-740.

Sutherland, S. 1987. Why hermaphroditic plants produce many more flowers than fruits: Experimental tests with Agave mckelveyana. Evolution 41:750-759.

Swofford, D.L. and R.B. Selander. 1981. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Hered. 72:281-283.

Tabachnik, B.G. and L.S. Fidell. 1983. Using multivariate statistics. Harper and Row, Publishers (New York). 509 pp.

Tanksley, S.D., H. Mdeina-Filho and C.M. Rick. 1981. The effect of isozyme selection on metric characters in an interspecific backcross of tomato: Basis of an early screening procedure. Theor. Appl. Genet. 60:291-296.

Thompson, P.A. 1981. Variations in seed size within populations of Silene dioca (L.) in relation to habitat. Ann. Bot. 47:623-634.

Turesson,G. 1922. The genotypical responses of the plant species to the habitat. Hereditas 3:211-350.

Vallejos, C.E. 1983. Enzyme activity staining. IN: Isozymes in plant genetics and breeding, Part A. (S.D. Tanksley and T.J. Orton, eds.) Elsevier Science Publishers (New York) pp. 469-516.

Van Schaik, P.H. and A.H. Probst. 1958a. The inheritance of inflorescence type, peduncle length, flowers per node and percent flower shedding in soybeans. Agron. J. 50:98-102.

Van Schaik, P.H. and A.H. Probst. 1958b. Effects of some environmental factors on flower production and reproductive efficiency in soybeans. Agron. J. 50: 192-197.

Vickery, R.K. Jr. and S.S. Hsu. 1984. Esterase variation associated with elevations, latitudes and ploidy level in populations of Mimulus glabratus Complex. Amer. Mid. Nat. 111:96-104.

Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. Amer. Stat. Assoc. J. 58:236-244.

Warwick, S.I., B.K. Thompson and L.D. Black. 1987. Genetic variation in Canadian and European populations of the colonizing weed species Apera spica-vent. New Phytol. 106:301-317.

Welsh, A.H., A.T. Peterson and S.A. Altmann. 1988. The fallacy of averages. Amer. Nat. 132:277-288.

Whightman, F. and J.C. Forrest. 1978. Properties of plant aminotransferases. Phytochem. 17:1455-1471.

Weibold, W.J., D.A. Ashley and H.R. Boerma. 1981. Reproductive abscission levels and patterns for eleven determinate soybean cultivars. Agron. J. 73:43-46.

Weil, R.R. and A.J. Ohlrogge. 1976. Components of soybean yield as influenced by canopy level and interplant competition. Agron. J. 68: 583-587.

Wetton, J.H., R.E. Carter, D.T. Parkin and D. Walters. 1987. Demographic study of a wild house sparrow population by DNA fingerprinting. Nature 327:147-149.

Williams, J.T. 1982. Genetic conservation of wild plants. Nature and Resources 18:14-15.

Wishart, D. 1987. CLUSTAN 3.2 User's Manual. Edinburgh University Press (Edinburgh). 250 pp.

Workman, P.L. and J.D. Niswander. 1970. Population studies on Southwestern Indian tribes. II. Local genetic differentiation in the Papago. Amer. J. Human Genet. 22:24-49.

Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323-354.

Wright, S. 1964. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19:395-420.

Wright, S. 1969. Evolution and the genetics of populations. Vol. 2. The theory of gene frequencies. University of Chicago Press (Chicago).

Wright, S. 1978. Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. University of Chicago Press (Chicago).

Wu, K.K. and S.K. Jain. 1978. Genetic and plastic responses in geographic differentiation of Bromus rubens populations. Can. J. Bot. 56:873-879.

Yamamoto, K. and Y. Nagato. 1984. Variation of DNA content in the genus Glycine. Jap. J. Breeding 34:163-170

Zhi-ang, H. and W. Hong-xin. 1985. Genetic structure of natural populations of wild soybean (Glycine soja) in Beijing region. Acta Botanica Sinica 27:599-604.

# APPENDICES

## Appendix I

### Descriptions of soybean seed collection sites (Mishima, Japan)

A. 1982 Collection

SAMPLE: A — COLLECTOR(S): H.I. Oka and H. Morishima
NO. OF PLANTS: 13 — DATE OF COLLECTION: 10/30/82

SITE LOCATION: Hatsunedai
SITE DESCRIPTION: A roadside strip along a rice field (unused land).
SITE DIMENSIONS: 5 X 1 meters
TOPOGRAPHY: Flat DISTURBANCE: Moderate COVER: 30%, clump

SAMPLE: B — COLLECTOR(S): H.I. Oka and H. Morishima
NO. OF PLANTS: 36 — DATE OF COLLECTION: 10/30/82

SITE LOCATION: Kanodanchi
SITE DESCRIPTION: A roadside strip used for ornamental plants (Rhododendron spp.).
SITE DIMENSIONS: 100 X 2.5 meters
TOPOGRAPHY: Flat DISTURBANCE: Moderate COVER: 5%, scattered

SAMPLE: C — COLLECTOR(S): H.I. Oka and H. Morishima
NO. OF PLANTS: 10 — DATE OF COLLECTION: 11/2/82

SITE LOCATION: Nishi-asahigaoka
SITE DESCRIPTION: A strip of land between the road and shrubs.
SITE DIMENSIONS: 30 x 2 meters
TOPOGRAPHY: Slope DISTURBANCE: Moderate COVER: 30-40%, clump

SAMPLE: D — COLLECTOR(S): H.I. Oka and H. Morishima
NO. OF PLANTS: 30 — DATE OF COLLECTION: 11/9/82

SITE LOCATION: Yata
SITE DESCRIPTION: An abandoned field on the campus of the National Institute of Genetics.
SITE DIMENSIONS: 10 X 1 meter
TOPOGRAPHY: Flat DISTURBANCE: High COVER: 60%, clump

Appendix I (continued)

Collectors comments to the 1982 wild soybean collection:

The wild soybeans climb up neighboring plants, so estimation of coverage is difficult. The sites are subject to accidental disturbance. Conditions are expected to change from year to year. Soil is sandy loam, brown to reddish brown, in all locations. The sites are well-drained, but the soil retains moisture if it rains intermittently.

B. 1983 Collection

SAMPLE: A COLLECTOR: H.I. Oka
NO. OF PLANTS: 10 DATE OF COLLECTION: 1983

SITE LOCATION: Asahigaoka-1
SITE DESCRIPTION: A ruderal field left for housing. Soil is sandy with gravel. Site is well drained and dry.
SITE DIMENSIONS: Not recorded (NR)
TOPOGRAPHY: NR DISTURBANCE: NR COVER: 10%

SAMPLE: B COLLECTOR: H.I. Oka
NO. OF PLANTS: 9 DATE OF COLLECTION: 1983

SITE LOCATION: Asahigaoka-2
SITE DESCRIPTION: A roadside slope. Dry. North face shaded.
SITE DIMENSIONS: Not recorded
TOPOGRAPHY: Slope DISTURBANCE: NR COVER: 20%

SAMPLE: C COLLECTOR: H.I. Oka
NO. OF PLANTS: 3 DATE OF COLLECTION: 1983

SITE LOCATION: Kakitagawa
SITE DESCRIPTION: A moist site on a river terrace. 75% shaded by trees.
SITE DIMENSIONS: NR
TOPOGRAPHY: Slope DISTURBANCE: NR COVER: 80%

Collectors comments to the 1983 wild soybean collection:

Wild soybeans climb over other plants. The soybeans seem to tolerate shading, but cannot tolerate drought when in competition with other species.

## Appendix II

### Summary of environmental variables among soybean seed collection sites

| Population | Environmental variables[a] | | | |
|---|---|---|---|---|
| | TOP[b] | SHADE | DIST | COVER |
| 1. Hatsunedai | 0 | 1 | 1 | 1 |
| 2. Kanodanchi | 0 | 1 | 1 | 0 |
| 3. Nishi-asahigaoka | 1 | 1 | 1 | 1 |
| 4. Yata | 0 | 1 | 2 | 2 |
| 5. Asahigaoka-1 | . | 1 | . | 0 |
| 6. Asahigaoka-2 | 1 | 0 | . | 0 |
| 7. Kakitagawa | 1 | 0 | . | 2 |

[a] Based on collector's notes (see Appendix I)
[b] Key:

TOP = topography 0=flat 1=slope

SHADE = shading 0=shading 1=no shading

DIST= site disturbance 0=low 1=moderate 2=high

COVER = % of collection site occupied by wild soybean relative to other co-habitating plant species 0=< 30% 1=30-50% 2=>50%

. = missing information

# Appendix III

## Gel making and enzyme-activity staining protocols

Abbreviations used:
APS.....Ammonium persulfate
NAD.....Nicotinamide adenine dinucleotide
NADH.....Nicotinamide adenine dinucleotide (reduced)
NADP.....Nicotinamide adenine dinucleotide phosphate
MTT.....4,5-dimethylthiozoyl-2)-2,5 diphenyltetrazolium bromide
PMS.....Phenazine methosulfate
TEMED.....N,N,N',N'-Tetramethylethylenediamine
TRIS.....Tris(hydroxymethyl)aminomethane
$Na_2$EDTA.....Ethylenediaminetetraacetate (disodium salt)

### I. Gel and Electrode Buffer Formulations

A. Gel Buffer: 0.005 M L-Histidine (pH 7.0)

| | | |
|---|---|---|
| L-Histidine (HCl) | 1.048 | gm |
| 4 M NaOH | 1.1 | ml |
| Distilled water | 1.0 | liter |

Chemicals were dissolved in distilled water and the pH was adjusted with 4 M NaOH.

B. Electrode Buffer: 0.13 M Tris-Citrate (pH 7.0)

| | | |
|---|---|---|
| Tris | 15.73 | gm |
| Citric acid | 7.68 | gm |
| Distilled water | 1.0 | liter |

Chemicals were dissolved in distilled water and the pH was adjusted with conc. HCl.

### II. Gel Types

This section summarizes the methods for single layer gel preparation. Two and three layer gels are made by modifying the amounts of chemicals and buffers as described in Chapter II, Materials and Methods.

a) <u>12.5% (w/v) starch</u>.

| | |
|---|---|
| 0.005 M Histidine (HCl) | 240 ml |
| Hydrolyzed potato starch (Sigma Electrophoresis Grade) | 30 gm |

Appendix III (continued)

The starch and buffer were combined in a 1000 ml Erlenmeyer side-arm flask containing a large magnetic stirring bar. The flask then was stoppered and placed in a hot water bath on a magnetic stirrer/hot plate. The starch solution was heated with constant stirring to 80° C. Any cofactors (e.g., NADP, NAD, etc.) then were added and allowed to mix into the heated starch solution. The flask then was degassed via vacuum aspiration for 40 seconds and poured immediately into a 2-layer gel mold.(NOTE: Due to the often brittle nature of starch gels, they were not made in single layer gel molds) A piece of plate glass (20 X 18 X 0.6 cm) was placed onto the gel to ensure an even surface.

b) 7% (w/v) acrylamide.

| | | |
|---|---|---|
| 0.005 M L-Histidine (HCl) | 150 | ml |
| Acrylamide | 9.975 | gm |
| N,N'-methylene-bis-acrylamide | 0.525 | gm |
| APS | 0.15 | gm |
| TEMED | 0.30 | ml |

All the chemicals, except TEMED, were added to the gel buffer in a 1000 ml glass beaker, mixed well, and heated to 30° C with constant stirring. TEMED then was added to the mixture and the solution immediately was poured into a gel mold.

c) 9% (w/v) acrylamide.

| | | |
|---|---|---|
| 0.005 M L-Histidine (HCl) | 150 | ml |
| Acrylamide | 12.82 | gm |
| N,N'-bis-acrylamide | 0.675 | gm |
| APS | 0.15 | gm |
| TEMED | 0.30 | ml |

The method for preparing this gel was the same as for b).

d) 7% (w/v) acrylamide + 2% (w/v) starch.

| | | |
|---|---|---|
| 0.005 M L-Histidine (HCl) | 150 | ml |
| Acrylamide | 9.975 | gm |
| N,N'-bis-acrylamide | 0.525 | gm |
| APS | 0.15 | gm |
| TEMED | 0.30 | ml |
| Hydrolyzed potato starch | 3.0 | gm |

The acrylamide chemicals, APS and any cofactors were added to 1/2 of the total volume of gel buffer and mixed well with constant stirring at room temperature. The starch and remaining volume of gel buffer were combined in an

Appendix III (continued)

Erlenmeyer side-arm flask, mixed well, stoppered and heated (with constant stirring) to 80° C. The solution then was degassed for 40 seconds. The hot starch solution was added to the acrylamide solution and allowed to mix thoroughly. TEMED then was added and the solution was poured immediately into a gel mold.

e) 6% (w/v) acrylamide + 4% (w/v) starch.

| | |
|---|---|
| 0.005 M L-Histidine (HCl) | 150 ml |
| Acrylamide | 8.55 gm |
| N,N'-methylene-bis-acyrlamide | 0.45 gm |
| APS | 0.15 gm |
| TEMED | 0.30 ml |
| Hydrolyzed potato starch | 6.0 gm |

The preparation of this gel was the same as described for d).

Gels were allowed to cool to room temperature. For gels b-e, any unpolymerized solution on the surface then was poured off. The gels were wrapped in plastic wrap and refrigerated at least eight hours before use.

III. Stain Buffer Formulations

The volume of distilled water in the following formulations is the final volume of the solution.

A. 0.2 M Tris-HCl (pH 8.5)

| | |
|---|---|
| Tris | 24.22 gm |
| Distilled water | 1000 ml |

Tris was dissolved in distilled water and the pH was adjusted with concentrated HCl.

B. 0.1 M Acetate Buffer (pH 5.0)

A. 0.2 M Acetic acid (11.55 ml in 1000 ml dHOH)
B. 0.2 M Sodium acetate (16.4 gm $C_2H_3O_2$ Na in 1000 ml dHOH)

148 ml A + 352 ml B were brought up to 1000 ml with dHOH.

Appendix III (continued)

C. 0.1 M Phosphate Buffer (pH 6.4)

| | | |
|---|---|---|
| Sodium monophosphate | 13.9 | gm |
| Sodium diphosphate | 5.3 | gm |
| Distilled water | 1000 | ml |

Chemicals were dissolved in distilled water and the pH was adjusted with conc. HCl.

D. 0.1 M Tris-Maleate Buffer (pH 5.2)

| | | |
|---|---|---|
| Tris | 12.1 | gm |
| Maleic acid | 11.6 | gm |
| Distilled water | 1000 | ml |

Chemicals were dissolved in distilled water and the pH was adjusted with 0.2 M HCl.

E. 0.2 M Tris-Maleate Buffer (pH 3.7)

| | | |
|---|---|---|
| Tris | 24.2 | gm |
| Maleic acid | 23.2 | gm |
| Distilled water | 1000 | ml |

Chemicals were dissolved in distilled water. For the enzyme assay, the stock was diluted as follows: 5(stock): 3(dHOH): 2(0.2 M NaOH). Final pH = 5.55.

F. 0.01 M Malate Buffer (pH 7.0)

| | | |
|---|---|---|
| L-Malic acid | 1.341 | gm |
| Tris | 3.025 | gm |
| Distilled water | 1000 | ml |

Chemicals were dissolved in distilled water and the pH was adjusted with 4 M NaOH.

G. 0.05 M Sodium Phosphate Buffer (pH 7.0)

A. 0.05 M Monobasic sodium phosphate
(6.9 gm $NaH_2PO_4$ in 1 liter dHOH)
B. 0.05 M Dibasic sodium phosphate
(7.1 gm $Na_2H_2PO_4$ in 1 liter dHOH)

39 mls A + 61 ml B. Adjust pH with 1 N HCl.

H. 0.05% Aniline Blue

| | | |
|---|---|---|
| Aniline blue | 0.05 | gm |
| 7% Acetic acid | 100 | ml |

The aniline blue was dissolved in the 7% acetic acid.

Appendix III (continued)

I. Potassium Iodide Solution

| | |
|---|---|
| Iodine | 0.10 gm |
| Potassium iodide | 0.50 gm |
| Distilled water | 100 ml |

The chemicals were dissolved in distilled water in an amber bottle and stirred slowly for several hours until the iodine had completely dissolved.

IV. Enzyme-activity stain protocols

Times given below for gel electrophoresis and enzyme activity staining are approximate, actual times depend on the thickness of the gel and on the ambient temperature.

A. OXIOREDUCTASES

Alcohol Dehydrogenase (ADH) EC 1.1.1.1

Gel: D + 15 mg NAD
Time: 8 hours at 200 volts

| | |
|---|---|
| 0.05 M Sodium phosphate buffer | 50 ml |
| MTT | 15 mg |
| NAD | 15 mg |
| 95% Ethanol | 5 ml |
| PMS | 2 mg |

The gel was placed in the dark at room temperature for 1 hour. Bands appeared blue against a white background.

Diaphorase (DIA) EC 1.6.2.2

Gel: D
Time: 12 hours at 200 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NADH | 10 mg |
| 2,6 Dichlorophenol indophenol | 2 mg |

The gel was incubated at 37°C in the dark for 2 hours. Bands appeared dark blue against a light blue background.

Appendix III (continued)

**Glucose 6-phosphate dehydrogenase (GPD) EC 1.1.1.49**

Gel: E + 15 mg NADP
Time: 24 hours at 150 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NADP | 15 mg |
| $Na_2$ Glucose 6-phosphate | 100 mg |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 3 hours. Bands appeared blue against a white background.

**Isocitrate dehyrodgenase (IDH) EC 1.1.1.42**

Gel: D
Time: 12 hours at 200 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NADP | 10 mg |
| $MgCl_2$ | 120 mg |
| DL-Isocitric acid | 200 mg |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 3 hours. Bands appeared blue against a white background.

**Malate dehydrogenase (MDH) EC 1.1.1.37**

Gel: D + 15 mg NAD
Time: 8 hours at 200 volts.

| | |
|---|---|
| 0.01 M Malate buffer | 50 ml |
| MTT | 15 mg |
| NAD | 20 mg |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 1 hour. Bands appeared blue against a white background.

Appendix III (continued)

**Peroxidase (PER)** EC 1.11.1.7

Gel: D
Time: 8 hours at 200 volts

| | |
|---|---|
| 3-Amino-9-ethylcarbazole (dissolved in 1.25 ml dimethyl formamide) | 25 mg |
| $CaCl_2$ | 1 ml |
| $H_2O_2$ (3% solution) | 1 ml |
| Sodium acetate | 3 ml |
| Distilled water (chilled) | 45 ml |

Chemicals were added slowly to prevent reprecipitation of 3-amino-9-ethylcarbazole. Gel was incubated in the dark at room temperature for 1 hour. Bands appeared red or brown against a white background.

**6-Phosphogluconate dehydrogenase (PGD)** EC 1.1.1.43

Gel: E
Time: 16 hours at 150 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NADP | 10 mg |
| $MgCl_2$ | 20 mg |
| 6-Phosphogluconic acid | 12 mg |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 1 hour. Bands appeared blue against a white background.

**Shikimate dehydrogenase (SKD)** EC 1.1.1.25

Gel: E + 15 mg NADP
Time: 24 hours at 150 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 15 mg |
| NADP | 15 mg |
| Shikimic acid | 15 mg |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 1 hour. Bands appeared blue against a white background.

Appendix III (continued)

B. TRANSFERASES

Glutamate oxaloacetic transaminase (GOT) EC 2.6.1.1

Gel: D
Time: 13 hours at 200 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| Pyridoxal 5'-phosphate | 25 mg |
| L-Aspartic acid | 266 mg |
| Alpha-ketoglutaric acid | 36 mg |
| Fast Blue BB salt | 112 mg |

The gel was incubated at 37°C in the dark for 1 hour. Bands appeared dark brown against a light brown background.

Phosphoglucomutase (PGM) EC 2.7.5.3

Gel: D
Time: 8 hours at 200 volts

| | |
|---|---|
| 0.02 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NAD | 10 mg |
| $MgCl_2$ | 20 mg |
| $Na_2$ D-Glucose 1-phosphate | 125 mg |
| Glucose 6-phosphate dehydrogenase (NAD active) | 40 units |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 2 hours. Bands appeared blue against a white background.

C. HYDROLASES

Acid phosphatase (AP) EC 3.1.3.2

Gel: D
Time: 12 hours at 200 volts

| | |
|---|---|
| 0.2M Acetate buffer | 50 ml |
| Black K salt | 40 mg |
| Alpha-napthyl acid phosphate | 40 mg |

The gel was placed at room temperature. Bands appeared dark brown against a light brown background.

Appendix III (continued)

**Beta-Amylase (AM)** EC 3.2.1.2

Gel: B
Time: 4 hours at 200 volts

| | |
|---|---|
| A. 0.2 M Acetate buffer | 100 ml |
| Soluble potato starch | 1 gm |

The potato starch was heated in acetate buffer until it dissolved and the solution turned clear. The solution then was cooled to 30°C and poured over the gel. The gel was incubated at 37°C in the dark for 15 to 30 minutes. Solution A then was poured off and the gel was rinsed well with distilled water.

B. 0.1% iodine in 0.5% potassium iodine

Twenty ml of B was poured over the gel. Bands appeared as clear areas against a blue background and were scored immediately.

**Endopeptidase (ENP)** EC 3.4.?.?

Gel: D
Time: 12 hours at 200 volts

| | |
|---|---|
| Endopeptidase buffer (0.2 M Tris-Maleate pH 3.7) | 50 ml |
| Black K salt | 20 mg |
| $MgCl_2$ | 10 mg |
| N-Alpha-benzoyl DL-arginine beta-naphthylamide (HCl) (BANA) | 20 mg |

Stain solution was stirred vigorously for 10 minutes in the dark then poured over the gel. The gel was placed at room temperature in the dark for 2 hours. Bands appeared brown against a light brown background.

**Esterase (EST)** EC 3.1.1.1

Gel: E
Time: 6 hours at 200 volts

| | | |
|---|---|---|
| 0.1 M Phosphate buffer | 50 | ml |
| Fast Blue RR salt | 80 | mg |
| Alpha-naphthyl butyrate* | 0.05 | cc |
| 100% Acetone | 3 | drops |
| 1% Alpha-naphthyl acetate | 1 | ml |

Appendix III (continued)

* This chemical comes in a sealed ampule and was delivered into the stain solution using a tuberculin syringe.

Alpha-naphthyl butyrate and three drops of acetone were added to the Fast Blue RR salt. Then 50 ml of phosphate buffer was added and the solution was stirred vigorously over low heat. One ml of 1% alpha-naphthyl acetate was added to the warm solution. The solution then was poured immediately over the gel through a single layer of cheesecloth. The gel was incubated at 37° C in the dark for 2 hours. Bands appeared grey or black against a light brown background.

Fluorescent esterase (FLE) EC 3.1.1.2

Gel: D
Time: 10 hours at 150 volts

| | |
|---|---|
| 0.2 M Acetate buffer | 50 ml |
| 100% Acetone | 10 ml |
| 4-Methylumbelleferyl acetate | 15 mg |

Fifteen mg of 4-methylumbelliferyl acetate was dissolved in acetone. To this solution was added 30 ml of acetate buffer. Three Kim-wipes were saturated with the stain solution and placed in direct contact with the sliced gel surface. After 5 minutes, the gel was examined under UV light (330nm) using a hand-held UV light source. Bands fluoresced under the UV light.

Leucine aminopeptidase (LAP) EC 3.4.11.1

Gel: D
Time: 8 hours at 200 volts

| | | |
|---|---|---|
| A. | 0.1 M Tris-Maleate buffer (25 mls stock + 75 mls dHOH) | 100 ml |
| | L-Leucine beta-naphthylamide (dissolved in 3 drops 100% acetone) | 20 mg |

The gel was incubated in A at 37°C for one hour. Solution A then was poured into a beaker.

| | | |
|---|---|---|
| B. | Black K salt | 50 mg |

Appendix III (continued)

B was added to A and poured back over gel. The gel was placed at room temperature in the dark for 1 hour. Bands appeared brown against a light brown background.

Urease (EU) EC 3.5.1.5

Gel: D
Time: 8 hours at 200 volts

| | | |
|---|---|---|
| A. | 0.2M Acetate buffer | 100 ml |
| | 0.1% (w/v) Cresol red | 0.1 gm |

Solution A was poured over the gel and the gel was placed at room temperature in the dark for 10-15 minutes. Solution A then was poured off.

| | | |
|---|---|---|
| B. | 333 mM Urea | 2 gm |
| | 0.1% (w/v) Cresol red | 0.1 gm |
| | 0.1% (w/v) $Na_2$EDTA | 0.1 gm |
| | Distilled water | 100 ml |

Solution B was poured over the gel. The gel was placed at room temperature in the dark. Bands appeared red on a white background within 15 minutes. Results were recorded immediately as the bands blurred within a few hours.

D. LYASES

Aconitase (ACO) EC 4.2.1.3

Gel: A + 15 mg NADP
Time: 6 hours at 160 volts

| | |
|---|---|
| 0.2M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NADP | 10 mg |
| $MgCl_2$ | 10 mg |
| cis-Aconitic Acid (1% solution, pH 7.5) | 8 ml |
| Isocitrate dehydrogenase | 40 units |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 3 hours. Bands appeared blue agains a white background.

Appendix III (continued)

E. ISOMERASES

Mannose 6-phosphate isomerase (MPI) EC 5.3.1.8

Gel: D
Time: 8 hours at 200 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NAD | 15 mg |
| Mannose 6-phosphate | 20 mg |
| Glucose 6-phosphate dehydrogenase (NAD active) | 40 units |
| Phosphoglucose isomerase | 40 units |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 1 hour. Bands appeared blue against a white background.

Phosphoglucose isomerase (PGI) EC 5.3.1.9

Gel: D
Time: 8 hours at 200 volts

| | |
|---|---|
| 0.2 M Tris-HCl | 50 ml |
| MTT | 10 mg |
| NAD | 10 mg |
| $MgCl_2$ | 20 mg |
| Fructose 6-phosphate | 30 mg |
| Glucose 6-phosphate dehydrogenase | 40 units |
| PMS | 1 mg |

The gel was incubated at 37°C in the dark for 1 hour. Bands appeared blue against a white background.

F. PROTEIN

Kunitz Trypsin inhibitor (Ti)

Gel: C
Time: 4 hours at 200 volts

A. 0.05 % Aniline blue solution

Aniline blue was poured over the gel. When dark bands appeared at the sample wells, the solution was poured

Appendix III (continued)

back into its original container (this solution can be reused several times).

B. 7% Acetic acid

Acetic acid was poured over the gel and replaced as needed. The gel was allowed to destain overnight. Bands appeared dark blue against a light blue background.

## Appendix IV

### Genotype counts for 48 isozyme loci among seven natural populations of wild soybean

| Population[a] | Ap | | Aco1 | Aco2 | | | Aco3 | Aco4 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | AA | CC | BB | AA | AB | BB | AA | AA | BB | CC |
| 1 | 0 | 13 | 13 | 10 | 0 | 3 | 13 | 10 | 0 | 3 |
| 2 | 0 | 36 | 36 | 34 | 1 | 1 | 36 | 1 | 0 | 35 |
| 3 | 0 | 10 | 10 | 10 | 0 | 0 | 10 | 0 | 0 | 10 |
| 4 | 0 | 30 | 30 | 30 | 0 | 0 | 30 | 0 | 30 | 0 |
| 5 | 1 | 9 | 10 | 8 | 1 | 1 | 10 | 0 | 0 | 10 |
| 6 | 0 | 9 | 9 | 8 | 0 | 1 | 9 | 0 | 0 | 9 |
| 7 | 0 | 3 | 3 | 3 | 0 | 0 | 3 | 0 | 0 | 3 |

| Population | Aco5 | Adh1 | Adh2[*] | Adh3 | Am3 | | |
|---|---|---|---|---|---|---|---|
| | AA | Adh1_ | XX | Adh3_ | AA | AB | BB |
| 1 | 13 | 13 | 13 | 13 | 0 | 0 | 13 |
| 2 | 36 | 36 | 36 | 36 | 27 | 2 | 7 |
| 3 | 10 | 10 | 10 | 10 | 10 | 0 | 0 |
| 4 | 30 | 30 | 30 | 30 | 30 | 0 | 0 |
| 5 | 10 | 10 | 10 | 10 | 8 | 0 | 2 |
| 6 | 9 | 9 | 9 | 9 | 6 | 1 | 2 |
| 7 | 3 | 3 | 3 | 3 | 3 | 0 | 0 |

| Population | Dia1 | | Dia2 | | Dia3 | Dia4 | Dia5[*] |
|---|---|---|---|---|---|---|---|
| | Dia1_ | dia1/dia1 | AA | BB | BB | AA | XX |
| 1 | 10 | 3 | 3 | 10 | 13 | 13 | 13 |
| 2 | 34 | 2 | 24 | 12 | 36 | 36 | 36 |
| 3 | 1 | 9 | 10 | 0 | 10 | 10 | 10 |
| 4 | 0 | 30 | 30 | 0 | 30 | 30 | 30 |
| 5 | 7 | 3 | 8 | 2 | 10 | 10 | 10 |
| 6 | 8 | 1 | 5 | 4 | 9 | 9 | 9 |
| 7 | 2 | 1 | 3 | 0 | 3 | 3 | 3 |

| Population | Enp | | Est1 | | | Fle | Gpd | |
|---|---|---|---|---|---|---|---|---|
| | AA | BB | AA | AB | BB | Fle_ | Gpd_ | gpd/gpd |
| 1 | 10 | 3 | 3 | 0 | 10 | 13 | 9 | 4 |
| 2 | 35 | 1 | 31 | 0 | 5 | 36 | 6 | 21 |
| 3 | 0 | 10 | 0 | 0 | 10 | 10 | 9 | 1 |
| 4 | 0 | 30 | 0 | 0 | 30 | 30 | 30 | 0 |
| 5 | 8 | 2 | 5 | 0 | 5 | 10 | 4 | 6 |
| 6 | 8 | 1 | 7 | 1 | 1 | 9 | 2 | 7 |
| 7 | 2 | 1 | 2 | 0 | 1 | 3 | 1 | 2 |

[a] Populations: 1=Hatsunedai; 2=Kanodanchi; 3=Nishi-asahigaoka; 4=Yata; 5=Asahigaoka-1; 6=Asahigaoka-2; 7=Kakitagawa.

[*] Hypothesized locus. All hypothesized loci were assumed to be homozygous.

Appendix IV (continued)

| Population | Got1 | Got2* | Got3* | Idh1 | Idh2 | | Idh3 | Idh4 |
|---|---|---|---|---|---|---|---|---|
| | BB | AA | AA | BB | AA | BB | BB | AA |
| 1 | 13 | 13 | 13 | 13 | 0 | 13 | 13 | 13 |
| 2 | 36 | 36 | 36 | 36 | 0 | 36 | 36 | 36 |
| 3 | 10 | 10 | 10 | 10 | 0 | 10 | 10 | 10 |
| 4 | 30 | 30 | 30 | 30 | 30 | 0 | 30 | 30 |
| 5 | 10 | 10 | 10 | 10 | 1 | 9 | 10 | 10 |
| 6 | 9 | 9 | 9 | 9 | 1 | 8 | 9 | 9 |
| 7 | 3 | 3 | 3 | 3 | 0 | 3 | 3 | 3 |

| Population | Ti | Lap | | Mdh1* | Mdh2* | Mdh3* | Mdh4* |
|---|---|---|---|---|---|---|---|
| | AA | AA | BB | XX | XX | XX | XX |
| 1 | 13 | 3 | 10 | 13 | 13 | 13 | 13 |
| 2 | 36 | 28 | 7 | 36 | 36 | 36 | 36 |
| 3 | 10 | 10 | 0 | 10 | 10 | 10 | 10 |
| 4 | 30 | 30 | 0 | 30 | 30 | 30 | 30 |
| 5 | 10 | 8 | 2 | 10 | 10 | 10 | 10 |
| 6 | 9 | 6 | 3 | 9 | 9 | 9 | 9 |
| 7 | 3 | 3 | 0 | 3 | 3 | 3 | 3 |

| Population | Mpi | | | Per1* | Per2* | Per3* | Pgd1 | |
|---|---|---|---|---|---|---|---|---|
| | BB | BC | CC | XX | XX | XX | BB | CC |
| 1 | 13 | 0 | 0 | 13 | 13 | 13 | 10 | 3 |
| 2 | 36 | 0 | 0 | 36 | 36 | 36 | 36 | 0 |
| 3 | 10 | 0 | 0 | 10 | 10 | 10 | 10 | 0 |
| 4 | 30 | 0 | 0 | 30 | 30 | 30 | 30 | 0 |
| 5 | 8 | 1 | 1 | 10 | 10 | 10 | 8 | 2 |
| 6 | 8 | 0 | 1 | 9 | 9 | 9 | 6 | 3 |
| 7 | 3 | 0 | 0 | 3 | 3 | 3 | 3 | 0 |

| Population | Pgd2 | Pgd3 | | Pgi1 | Pgi2 | Pgi3 | Pgm1 | |
|---|---|---|---|---|---|---|---|---|
| | AA | AA | BB | AA | pgi2_ | BB | AA | BB |
| 1 | 13 | 0 | 13 | 13 | 13 | 13 | 13 | 0 |
| 2 | 36 | 0 | 36 | 36 | 36 | 36 | 36 | 0 |
| 3 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 0 |
| 4 | 30 | 30 | 0 | 30 | 30 | 30 | 0 | 30 |
| 5 | 10 | 0 | 10 | 10 | 10 | 10 | 10 | 0 |
| 6 | 9 | 0 | 9 | 9 | 9 | 9 | 9 | 0 |
| 7 | 3 | 0 | 3 | 3 | 3 | 3 | 3 | 0 |

| Population | Pgm2 | Pgm3 | Skd1* | Skd2* | Eu |
|---|---|---|---|---|---|
| | BB | Pgm3_ | XX | XX | AA |
| 1 | 13 | 13 | 13 | 13 | 13 |
| 2 | 36 | 36 | 36 | 36 | 36 |
| 3 | 10 | 10 | 10 | 10 | 10 |
| 4 | 30 | 30 | 30 | 30 | 30 |
| 5 | 10 | 10 | 10 | 10 | 10 |
| 6 | 9 | 9 | 9 | 9 | 9 |
| 7 | 3 | 3 | 3 | 3 | 3 |

## Appendix V

### Allele frequencies for 48 isozyme loci among seven natural populations of wild soybean

| | Population | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| | N (13) | (36) | (10) | (30) | (10) | (9) | (3) | |
| Locus | | | | | | | | Mean |
| Ap | | | | | | | | |
| a | 0 | 0 | 0 | 0 | 0.100 | 0 | 0 | 0.014 |
| c | 1.000 | 1.000 | 1.000 | 1.000 | 0.900 | 1.000 | 1.000 | 0.986 |
| Aco1 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Aco2 | | | | | | | | |
| a | 0.769 | 0.958 | 1.000 | 1.000 | 0.850 | 0.889 | 1.000 | 0.924 |
| b | 0.231 | 0.042 | 0 | 0 | 0.150 | 0.111 | 0 | 0.076 |
| Aco3 | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Aco4 | | | | | | | | |
| a | 0.231 | 0.028 | 0 | 0 | 0 | 0 | 0 | 0.037 |
| b | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.143 |
| c | 0.769 | 0.972 | 1.000 | 0 | 1.000 | 1.000 | 1.000 | 0.820 |
| Aco5 | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Adh1 | | | | | | | | |
| Adh1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Adh2* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Adh3 | | | | | | | | |
| Adh3 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Am3 | | | | | | | | |
| a(s) | 0 | 0.778 | 1.000 | 1.000 | 0.800 | 0.611 | 1.000 | 0.741 |
| b(f) | 1.000 | 0.222 | 0 | 0 | 0.200 | 0.389 | 0 | 0.259 |
| Dia1 | | | | | | | | |
| Dia1 | 0.231 | 0.056 | 0.900 | 1.000 | 0.300 | 0.111 | 0.333 | 0.419 |
| dia1 | 0.769 | 0.944 | 0.100 | 0 | 0.700 | 0.889 | 0.667 | 0.581 |

Appendix V (continued)

| Locus | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean |
|---|---|---|---|---|---|---|---|---|
| Dia2 | | | | | | | | |
| a | 0.231 | 0.667 | 1.000 | 1.000 | 0.800 | 0.556 | 1.000 | 0.751 |
| b | 0.769 | 0.333 | 0 | 0 | 0.200 | 0.444 | 0 | 0.249 |
| Dia3 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Dia4 | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Dia5* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Enp | | | | | | | | |
| a | 0.769 | 0.972 | 0 | 0 | 0.700 | 0.889 | 0.667 | 0.571 |
| b | 0.231 | 0.028 | 1.000 | 1.000 | 0.300 | 0.111 | 0.333 | 0.429 |
| Est1 | | | | | | | | |
| a | 0.231 | 0.861 | 0 | 0 | 0.500 | 0.833 | 0.667 | 0.442 |
| b | 0.769 | 0.139 | 1.000 | 1.000 | 0.500 | 0.167 | 0.333 | 0.558 |
| Fle | | | | | | | | |
| Fle | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Gpd | | | | | | | | |
| Gpd | 0.692 | 0.222 | 0.900 | 1.000 | 0.400 | 0.111 | 0.333 | 0.523 |
| gpd | 0.308 | 0.778 | 0.100 | 0 | 0.600 | 0.889 | 0.667 | 0.477 |
| Got1 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Got2* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Got3* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Idh1 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Idh2 | | | | | | | | |
| a | 0 | 0 | 0 | 1.000 | 0.100 | 0.111 | 0 | 0.173 |
| b | 1.000 | 1.000 | 1.000 | 0 | 0.900 | 0.889 | 1.000 | 0.827 |

Appendix V (continued)

| Locus | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean |
|---|---|---|---|---|---|---|---|---|
| Idh3 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Idh4 | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Ti | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Lap | | | | | | | | |
| a | 0.231 | 0.800 | 1.000 | 1.000 | 0.800 | 0.667 | 1.000 | 0.785 |
| b | 0.769 | 0.200 | 0 | 0 | 0.200 | 0.333 | 0 | 0.215 |
| Mdh1* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Mdh2* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Mdh3* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Mdh4* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Mpi | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 0.850 | 0.889 | 1.000 | 0.963 |
| c | 0 | 0 | 0 | 0 | 0.150 | 0.111 | 0 | 0.037 |
| Per1* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Per2* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Per3* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Pgd1 | | | | | | | | |
| b | 0.769 | 1.000 | 1.000 | 1.000 | 0.800 | 0.667 | 1.000 | 0.891 |
| c | 0.231 | 0 | 0 | 0 | 0.200 | 0.333 | 0 | 0.109 |
| Pgd2 | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Appendix V (continued)

| Locus | Population 1 | 2 | 3 | 4 | 5 | 6 | 7 | Mean |
|---|---|---|---|---|---|---|---|---|
| Pgd3 | | | | | | | | |
| a | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.143 |
| b | 1.000 | 1.000 | 1.000 | 0 | 1.000 | 1.000 | 1.000 | 0.857 |
| Pgi1 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Pgi2 | | | | | | | | |
| pgi2 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Pgi3 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Pgm1 | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 0 | 1.000 | 1.000 | 1.000 | 0.857 |
| b | 0 | 0 | 0 | 1.000 | 0 | 0 | 0 | 0.143 |
| Pgm2 | | | | | | | | |
| b | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Pgm3 | | | | | | | | |
| Pgm3 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Skd1* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Skd2* | | | | | | | | |
| x | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Eu | | | | | | | | |
| a | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

* Loci hypothesized based on zymogram patterns.

## Appendix VI

### Canonical discriminant analysis of electrophoretic data

A. Between canonical structure

| Allele | Canonical variate | | | |
|---|---|---|---|---|
| | CAN1 | CAN2 | CAN3 | CAN4 |
| Aco2-a | 0.515 | -0.606 | 0.334 | 0.472 |
| Aco2-b | -0.515 | 0.606 | -0.334 | -0.472 |
| Aco4-a | -0.379 | 0.645 | -0.651 | 0.079 |
| Aco4-b | 1.000 | 0 | 0 | 0 |
| Aco4-c | -0.994 | -0.074 | 0.074 | -0.009 |
| Am3-a (s) | 0.510 | -0.685 | 0.482 | 0.062 |
| Am3-b (f) | -0.510 | 0.685 | -0.482 | -0.062 |
| Ap-a | -0.213 | -0.006 | 0.345 | -0.875 |
| Ap-c | 0.213 | 0.006 | -0.345 | 0.875 |
| Dia1 (+) | 0.820 | -0.529 | -0.156 | -0.154 |
| dia1 (-) | -0.820 | 0.529 | 0.156 | 0.154 |
| Dia2-a | 0.454 | -0.207 | -0.439 | 0.261 |
| Dia2-b | -0.454 | 0.207 | 0.439 | -0.261 |
| Enp-a | -0.787 | 0.506 | 0.316 | 0.159 |
| Enp-b | 0.787 | -0.506 | -0.316 | -0.159 |
| Est-a | -0.697 | 0.255 | 0.664 | 0.041 |
| Est-b | 0.697 | -0.255 | -0.664 | -0.041 |
| Gpd | 0.759 | -0.232 | -0.606 | 0.044 |
| gpd | -0.759 | 0.232 | 0.606 | -0.044 |
| Idh2-a | 0.996 | 0.003 | 0.046 | -0.057 |
| Idh2-b | -0.996 | -0.003 | -0.046 | 0.057 |
| Lap-a | 0.543 | -0.690 | 0.431 | 0.084 |
| Lap-b | -0.543 | 0.690 | -0.431 | -0.084 |
| Mpi-b | 0.305 | -0.024 | -0.494 | 0.742 |
| Mpi-c | -0.305 | 0.024 | 0.494 | -0.742 |
| Pgd1-b | 0.422 | -0.405 | 0.064 | 0.500 |
| Pgd1-c | -0.422 | 0.405 | -0.064 | -0.500 |
| Pgd3-a | 1.000 | 0 | 0 | 0 |
| Pgd3-b | -1.000 | 0 | 0 | 0 |
| Pgm1-a | -1.000 | 0 | 0 | 0 |
| Pgm1-b | 1.000 | 0 | 0 | 0 |

Appendix VI (continued)

B. Standardized canonical coefficients

| Allele | Canonical variate | | | |
|---|---|---|---|---|
| | CAN1 | CAN2 | CAN3 | CAN4 |
| Aco2-a | 0 | -2.949 | 2.402 | 1.045 |
| Aco2-b | 0 | 0 | 0 | 0 |
| Aco4-a | 0 | -0.651 | 0.935 | 0.207 |
| Aco4-b | 3.918 | 6.955 | 5.501 | -9.051 |
| Aco4-c | 0 | 0 | 0 | 0 |
| Am3-a (s) | 0 | -2.396 | 0.772 | -1.212 |
| Am3-b (f) | 0 | 0 | 0 | 0 |
| Ap-a | 0 | 0.035 | 0.128 | -1.066 |
| Ap-c | 0 | 0 | 0 | 0 |
| Dia1 (+) | 0 | 0 | 0 | 0 |
| dia1 (-) | 0 | -2.181 | 0.825 | 2.041 |
| Dia2-a | 0 | 1.123 | -1.380 | -0.139 |
| Dia2-b | 0 | 0 | 0 | 0 |
| Enp-a | 0 | 4.943 | 0.080 | 0.139 |
| Enp-b | 0 | 0 | 0 | 0 |
| Est-a | 0 | -0.861 | 2.531 | 0.182 |
| Est-b | 0 | 0 | 0 | 0 |
| Gpd | 0 | 0 | 0 | 0 |
| gpd | 0 | 1.664 | -1.141 | -2.026 |
| Idh2-a | 0 | -3.894 | -4.377 | 9.209 |
| Idh2-b | 0 | 0 | 0 | 0 |
| Lap-a | 0 | 0.439 | 0.422 | 2.589 |
| Lap-b | 0 | 0 | 0 | 0 |
| Mpi-b | 0 | -0.134 | -3.030 | 0.918 |
| Mpi-c | 0 | 0 | 0 | 0 |
| Pgd1-b | 0 | -0.093 | 0.199 | 0.624 |
| Pgd1-c | 0 | 0 | 0 | 0 |
| Pgd3-a | 0 | 0 | 0 | 0 |
| Pgd3-b | 0 | 0 | 0 | 0 |
| Pgm1-a | 0 | 0 | 0 | 0 |
| Pgm1-b | 0 | 0 | 0 | 0 |

## Appendix VII

### Summary of quantitative traits used in univariate and multivariate analyses ('+'=included; '-'=excluded)

A. 1986 analysis

| Trait | | Univariate | | Multivariate | | |
|---|---|---|---|---|---|---|
| | | ANOVA | Spearman rank correlation | $D^2$ | CDA | Cluster analysis |
| 1. | DTG | + | + | + | + | + |
| 2. | DTF | + | + | + | + | + |
| 3. | DFP | + | + | + | + | + |
| 4. | DDP | + | + | + | + | + |
| 5. | LDP[b] | + | + | + | + | + |
| 6. | LSN | + | + | + | + | + |
| 7. | B4W[a] | + | + | + | + | + |
| 8. | L4W | + | + | + | + | + |
| 9. | H4W[a,g] | + | - | - | - | - |
| 10. | GH | + | + | + | + | + |
| 11. | L7W | + | + | + | + | + |
| 12. | ONO | + | + | + | + | + |
| 13. | BPW | + | + | + | + | + |
| 14. | FL[b] | + | + | + | + | + |
| 15. | CTL | + | + | + | + | + |
| 16. | LCL | + | + | + | + | + |
| 17. | WCL | + | + | + | + | + |
| 18. | LRC[f] | + | + | - | - | - |
| 19. | LLL | + | + | + | + | + |
| 20. | WLL[b] | + | + | + | + | + |
| 21. | LRL[f] | + | + | - | - | - |
| 22. | RF[d] | - | - | - | - | - |
| 23. | AGW[e] | + | + | - | - | - |
| 24. | BGW[e] | + | + | - | - | - |
| 25. | NDW[e] | + | + | - | - | - |
| 26. | NRN[e] | + | + | - | - | - |
| 27. | PPP | + | + | + | + | + |
| 28a. | P1SP[f] | + | + | - | - | - |
| 28b. | P2SP[f] | + | + | - | - | - |
| 28c. | P3SP[f] | + | + | - | - | - |
| 28d. | P4SP[f] | + | + | - | - | - |
| 29. | SWP | + | + | + | + | + |
| 30. | SWH[b] | + | + | + | + | + |
| 31. | SPP | + | + | + | + | + |
| 32. | ASP[a,f] | + | + | - | - | - |

Appendix VII (continued)

B. 1987 analysis

| Trait | | Univariate | | Multivariate | | |
|---|---|---|---|---|---|---|
| | | ANOVA | Spearman rank correlation | $D^2$ | CDA | Cluster analysis |
| 1. | DTG | + | + | + | + | + |
| 2. | DTF | + | + | + | + | + |
| 3. | DFP | + | + | + | + | + |
| 4. | DDP | + | + | + | + | + |
| 5. | LDP[b] | + | + | + | + | + |
| 6. | LSN | + | + | + | + | + |
| 7. | B4W | + | + | + | + | + |
| 8. | L4W[a] | + | + | + | + | + |
| 9. | H4W[b,c] | + | - | - | - | - |
| 10. | GH | + | + | + | + | + |
| 11. | L7W | + | + | + | + | + |
| 12. | ONO | + | + | + | + | + |
| 13. | BPW[a] | + | + | + | + | + |
| 14. | FL | + | + | + | + | + |
| 15. | CTL | + | + | + | + | + |
| 16. | LCL | + | + | + | + | + |
| 17. | WCL | + | + | + | + | + |
| 18. | LRC[b,f] | + | + | - | - | - |
| 19. | LLL | + | + | + | + | + |
| 20. | WLL | + | + | + | + | + |
| 21. | LRL[b,f] | + | + | - | - | - |
| 22. | RF[d] | - | - | - | - | - |
| 23. | AGW[e] | + | + | - | - | - |
| 24. | BGW[e] | + | + | - | - | - |
| 25. | NDW[a,e] | + | + | - | - | - |
| 26. | NRN[e] | + | + | - | - | - |
| 27. | PPP | + | + | + | + | + |
| 28a. | P1SP[f] | + | + | - | - | - |
| 28b. | P2SP[f] | + | + | - | - | - |
| 28c. | P3SP[f] | + | + | - | - | - |
| 28d. | P4SP[f] | + | + | - | - | - |
| 29. | SWP | + | + | + | + | + |
| 30. | SWH | + | + | + | + | + |
| 31. | SPP | + | + | + | + | + |
| 32. | ASP[f] | + | + | - | - | - |

[a] Square-root transformed variable.
[b] Log (base 10) transformed variable.
[c] Variable eliminated due to severe violations of model assumptions.
[d] Root fluorescence was the same for all plants tested. It was, therefore, eliminated from all statistical analyses.

Appendix VII (continued)

- Eliminated from multivariate analyses due to small number cases in the data set.
- Eliminated from multivariate analyses to prevent multicollinearity and/or singularity.
- Eliminated to make 1986 and 1987 multivariate analyses comparable.

## Appendix VIII

### Canonical discriminant analysis of quantitative traits

I. Between canonical structure

A. 1986 between canonical structure

| Character[a] | Canonical variate[b] CAN1 | CAN2 | CAN3 |
|---|---|---|---|
| 1. DTG | -0.179 | -0.670 | 0.296 |
| 2. DTF | 0.102 | 0.542 | 0.388 |
| 3. DFP | -0.105 | -0.623 | -0.349 |
| 4. DDP | 0.186 | -0.185 | 0.253 |
| 5. LDP | -0.876 | -0.469 | -0.022 |
| 6. LSN | -0.099 | 0.097 | 0.403 |
| 7. B4W | 0.543 | -0.400 | -0.328 |
| 8. L4W | 0.210 | -0.224 | 0.411 |
| 10. GH | -0.440 | -0.661 | 0.498 |
| 11. L7W | 0.071 | -0.561 | -0.186 |
| 12. ONO | 0.832 | -0.240 | 0.435 |
| 13. BPW | 0.769 | -0.625 | 0.133 |
| 14. FL | 0.935 | 0.257 | 0.065 |
| 15. CTL | 0.828 | 0.288 | -0.142 |
| 16. LCL | 0.319 | -0.256 | -0.695 |
| 17. WCL | 0.772 | 0.505 | -0.377 |
| 19. LLL | 0.691 | -0.047 | -0.593 |
| 20. WLL | 0.802 | 0.559 | 0.005 |
| 27. PPP | -0.671 | -0.669 | -0.095 |
| 29. SWP | 0.070 | -0.803 | 0.337 |
| 30. SWH | 0.933 | 0.268 | -0.198 |
| 31. SPP | -0.636 | -0.631 | 0.346 |

B. 1987 between canonical structure

| Character[a] | Canonical variate[b] CAN1 | CAN2 |
|---|---|---|
| 1. DTG | -0.344 | 0.322 |
| 2. DTF | 0.444 | 0.255 |
| 3. DFP | -0.089 | -0.369 |
| 4. DDP | -0.375 | -0.293 |
| 5. LDP | -0.775 | 0.058 |
| 6. LSN | 0.148 | 0.638 |
| 7. B4W | -0.188 | 0.683 |

Appendix VIII (continued)

B. 1987 Between canonical structure (continued)

| Character[a] | Canonical variate[b] CAN1 | CAN2 |
|---|---|---|
| 8. L4W | -0.520 | 0.539 |
| 10. GH | -0.613 | 0.405 |
| 11. L7W | -0.798 | 0.478 |
| 12. ONO | 0.343 | 0.675 |
| 13. BPW | -0.219 | 0.945 |
| 14. FL | -0.374 | 0.895 |
| 15. CTL | -0.203 | 0.768 |
| 16. LCL | -0.653 | -0.669 |
| 17. WCL | 0.920 | -0.160 |
| 19. LLL | -0.579 | 0.126 |
| 20. WLL | 0.696 | 0.158 |
| 27. PPP | -0.871 | -0.043 |
| 29. SWP | -0.419 | 0.292 |
| 30. SWH | 0.884 | -0.028 |
| 31. SPP | -0.887 | 0.199 |

II. Standardized canonical coefficients

A. 1986 standardized canonical coefficients

| Character[a] | Canonical variate[b] CAN1 | CAN2 | CAN3 |
|---|---|---|---|
| 1. DTG | -0.337 | -0.095 | 0.163 |
| 2. DTF | -0.044 | -0.556 | -0.353 |
| 3. DFP | -0.333 | -1.016 | -0.248 |
| 4. DDP | 0.360 | 0.388 | -0.287 |
| 5. LDP | 0.025 | 0.294 | 0.056 |
| 6. LSN | -0.677 | -0.518 | -0.067 |
| 7. B4W | 0.092 | 0.356 | 0.071 |
| 8. L4W | -0.352 | -0.034 | 0.293 |
| 10. GH | -0.234 | -0.696 | -0.025 |
| 11. L7W | 0.105 | -0.078 | 0.179 |
| 12. ONO | 0.243 | 0.259 | 0.023 |
| 13. BPW | 1.654 | -3.756 | 0.164 |
| 14. FL | 1.943 | 2.037 | 0.316 |
| 15. CTL | 0.497 | 0.794 | 0.026 |
| 16. LCL | -2.291 | 0.404 | 0.479 |
| 17. WCL | 0.452 | -0.449 | -1.553 |
| 19. LLL | 3.008 | -0.577 | -1.482 |

Appendix VIII (continued)

A. 1986 standardized canonical coefficients (cont)

| Character[a] | Canonical variate[b] CAN1 | CAN2 | CAN3 |
|---|---|---|---|
| 20. WLL | -1.415 | -0.146 | 1.701 |
| 27. PPP | -0.578 | -2.130 | -2.257 |
| 29. SWP | -3.471 | -1.912 | -0.335 |
| 30. SWH | 3.344 | 2.143 | 0.213 |
| 31. SPP | 4.992 | 4.166 | 2.462 |

B. 1987 standardized canonical coefficients

| Character[a] | Canonical variate[b] CAN1 | CAN2 |
|---|---|---|
| 1. DTG | 0.331 | -0.107 |
| 2. DTF | 3.079 | 0.347 |
| 3. DFP | 1.854 | -0.416 |
| 4. DDP | -0.327 | -0.274 |
| 5. LDP | -0.103 | -0.458 |
| 6. LSN | 0.032 | 0.485 |
| 7. B4W | 0.065 | 0.396 |
| 8. L4W | 0.338 | -0.721 |
| 10. GH | -0.797 | -0.473 |
| 11. L7W | -0.007 | -0.170 |
| 12. ONO | 0.585 | 0.027 |
| 13. BPW | 0.734 | 2.068 |
| 14. FL | 0.223 | -0.809 |
| 15. CTL | -0.939 | 0.344 |
| 16. LCL | -1.527 | -1.099 |
| 17. WCL | 6.570 | -1.762 |
| 19. LLL | 1.512 | 0.653 |
| 20. WLL | -6.097 | 0.947 |
| 27. PPP | 0.039 | -0.415 |
| 29. SWP | -2.618 | -0.697 |
| 30. SWH | 2.724 | 1.264 |
| 31. SPP | 4.551 | 2.094 |

[a] For the 1986 data set, three canonical functions were significant as indicated by Wilk's lambda; for 1987, two canonical functions were significant.

[b] For explanation of character notation see Chapter 3, Materials and Methods.

## Appendix IX

### Spearman rank correlation coefficients between environmental variables and population means for quantitative traits

A. 1986 Spearman rank coefficients

| | Environmental variables | | | |
|---|---|---|---|---|
| Character | TOP | SHADE | DIST | COVER |
| DTG | 0.30 | -0.66 | -0.54 | -0.61 |
| DTF | 0.48 | -0.16 | 0.26 | 0 |
| DFP | -0.09 | -0.32 | 0.26 | 0.17 |
| DDP | 0.29 | -0.79* | 0.26 | 0.23 |
| LDP | -0.10 | 0.16 | -0.82 | -0.49 |
| LSN | 0.68 | -0.79* | -0.26 | -0.38 |
| B4W | -0.68 | 0.47 | 0.26 | -0.07 |
| L4W | 0.10 | -0.24 | 0.27 | 0.12 |
| GH | 0.10 | -0.40 | -0.77 | -0.46 |
| L7W | -0.89* | 0.79* | -0.27 | -0.47 |
| ONO | -0.39 | 0.08 | 0.77 | -0.13 |
| BPW | -0.59 | 0.24 | 0.77 | -0.21 |
| FL | 0 | -0.24 | 0.82 | 0.30 |
| CTL | -0.44 | 0 | 1.00** | 0.64 |
| LCL | -0.49 | 0.63 | 0.26 | 0.11 |
| WCL | 0 | 0.16 | 0.77 | 0.50 |
| LRC | -0.29 | 0.24 | -0.77 | -0.63 |
| LLL | -0.49 | 0.47 | 0.77 | 0.09 |

Appendix IX (continued)

A. 1986 Spearman rank coefficients (continued)

| | Environmental variables | | | |
|---|---|---|---|---|
| Character | TOP | SHADE | DIST | COVER |
| WLL | 0.29 | -0.16 | 0.77 | 0.62 |
| LRL | -0.39 | 0.33 | -0.77 | -0.68 |
| AGW | 0.29 | -0.16 | 0.77 | 0.51 |
| BGW | -0.88* | 0.63 | 0.26 | -0.34 |
| NRN | 0.29 | -0.16 | 0.77 | 0.86 |
| NDW | 0.29 | 0.08 | 0.77 | 0.60 |
| PPP | -0.29 | 0.16 | -0.77 | -0.58 |
| P1SP | 0.10 | 0.48 | -0.26 | 0.26 |
| P2SP | -0.49 | 0.80* | -0.26 | 0.13 |
| P3SP | 0.10 | -0.63 | 0.26 | -0.13 |
| P4SP | -0.30 | 0 | 0.82 | 0.22 |
| SWP | -0.49 | 0.16 | -0.27 | -0.49 |
| SWH | -0.20 | 0.24 | 0.82 | -0.10 |
| SPP | -0.10 | 0 | -0.77 | -0.53 |
| ASP | 0.10 | -0.32 | 0.26 | 0.04 |

* and ** = product-moment coefficient significantly different from zero at 5% and 1% levels, respectively.

Appendix IX (continued)

B. 1987 Spearman rank coefficients

| | Environmental variables | | | |
|---|---|---|---|---|
| Character | TOP | SHADE | DIST | COVER |
| DTG | -0.10 | 0 | -0.33 | -0.91** |
| DTF | 0.10 | -0.47 | 0.26 | -0.13 |
| DFP | -0.29 | 0.63 | -0.26 | 0.13 |
| DDP | -0.29 | 0.63 | -0.26 | 0.13 |
| LDP | 0.20 | -0.08 | -0.26 | 0.57 |
| LSN | -0.39 | 0.40 | -0.77 | -0.67 |
| B4W | -0.29 | -0.16 | 0.77 | 0.52 |
| L4W | -0.10 | -0.44 | 0 | 0.10 |
| GH | 0.10 | -0.48 | -0.77 | -0.54 |
| L7W | 0.10 | -0.63 | -0.26 | 0.08 |
| ONO | -0.39 | 0.16 | 0.82 | -0.13 |
| BPW | -0.30 | -0.16 | 0.82 | -0.24 |
| FL | -0.30 | -0.16 | -0.27 | -0.66 |
| CTL | -0.73 | 0.33 | 0.26 | -0.31 |
| LCL | 0 | 0.40 | -0.77 | -0.38 |
| WCL | 0.29 | -0.07 | 0.78 | 0.59 |
| LRC | -0.29 | 0.16 | -0.77 | -0.59 |
| LLL | -0.40 | 0.32 | -0.54 | -0.72 |
| WLL | -0.29 | 0.08 | 0.77 | 0.10 |
| LRL | 0 | 0.16 | -0.82 | -0.62 |
| AGW | -0.68 | 0.16 | 0.77 | 0.08 |

Appendix IX (continued)

B. 1987 Spearman rank coefficients (continued)

| | Environmental variables | | | |
|---|---|---|---|---|
| Character | TOP | SHADE | DIST | COVER |
| BGW | -0.88* | 0.32 | 0.26 | -0.19 |
| NRN | -0.10 | 0.32 | -0.77 | 0.47 |
| NDW | -0.88* | 0.32 | 0.77 | 0.42 |
| PPP | -0.10 | 0.16 | -0.77 | -0.66 |
| P1SP | -0.10 | 0.63 | 0.26 | 0.45 |
| P2SP | -0.29 | 0.63 | -0.26 | 0.13 |
| P3SP | -0.10 | -0.47 | 0.26 | -0.47 |
| P4SP | -0.10 | -0.32 | 0.26 | -0.30 |
| SWP | -0.10 | -0.08 | 0.26 | -0.11 |
| SWH | -0.29 | 0.63 | 0.77 | 0.19 |
| SPP | 0.10 | -0.47 | -0.77 | -0.24 |
| ASP | 0.40 | -0.80* | 0.27 | -0.31 |

* and ** = product-moment coefficient significantly different from zero at 5% and 1% levels, respectively.

# Appendix X

## Squared Euclidean ($d^2$) distance matrices

A. Squared Euclidean distance ($d^2$) matrix for wild soybean isozyme data (Chapters 2 and 4)

| | Population | | | | | |
|---|---|---|---|---|---|---|
| Population | 1 | 2 | 3 | 4 | 5 | 6 |
| 2 | 1.674 | | | | | |
| 3 | 2.775 | 1.507 | | | | |
| 4 | 4.569 | 3.590 | 1.941 | | | |
| 5 | 2.086 | 1.160 | 1.808 | 3.658 | | |
| 6 | 1.371 | 0.523 | 2.183 | 4.002 | 0.766 | |
| 7 | 2.448 | 0.265 | 0.731 | 2.789 | 1.193 | 0.962 |

B. Squared Euclidean distance ($d^2$) matrix for wild soybean quantitative traits (1986 data) (Chapters 3 and 4)

| | Population | | | | | |
|---|---|---|---|---|---|---|
| Population | 1 | 2 | 3 | 4 | 5 | 6 |
| 2 | 1.644 | | | | | |
| 3 | 2.242 | 1.175 | | | | |
| 4 | 3.765 | 3.485 | 3.001 | | | |
| 5 | 1.368 | 1.097 | 2.268 | 3.214 | | |
| 6 | 2.029 | 0.613 | 1.347 | 2.990 | 1.246 | |
| 7 | 2.356 | 0.972 | 1.427 | 3.296 | 1.721 | 0.751 |

C. Squared Euclidean distance ($d^2$) matrix for wild soybean quantitative traits (1987 data) (Chapters 3 and 4)

| | Population | | | | | |
|---|---|---|---|---|---|---|
| Population | 1 | 2 | 3 | 4 | 5 | 6 |
| 2 | 2.657 | | | | | |
| 3 | 3.051 | 2.726 | | | | |
| 4 | 4.027 | 1.887 | 1.753 | | | |
| 5 | 2.012 | 0.725 | 1.919 | 1.234 | | |
| 6 | 2.386 | 0.848 | 2.432 | 1.999 | 0.567 | |
| 7 | 3.411 | 1.196 | 2.225 | 2.021 | 1.301 | 1.653 |

## Appendix XI

A. 1986 Spearman rank correlation coefficients between isozyme loci and quantitative traits

| | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| Locus | DTG | DTF | DFP | DDP | LDP | LSN |
| Ap-a | 0.21 | -0.41 | 0.20 | -0.61 | -0.21 | 0.20 |
| Ap-c | -0.21 | 0.41 | -0.20 | 0.61 | 0.21 | -0.20 |
| Aco2-a | -0.39 | 0.63 | -0.56 | 0.33 | -0.62 | 0.07 |
| Aco2-b | 0.39 | -0.63 | 0.56 | -0.33 | 0.62 | -0.07 |
| Aco4-a | -0.05 | -0.53 | 0.27 | -0.04 | 0.69 | -0.53 |
| Aco4-b | -0.53 | 0.20 | -0.20 | 0 | -0.64 | -0.41 |
| Aco4-c | 0.50 | 0.24 | -0.04 | 0.04 | 0.02 | 0.77* |
| Am3-s | -0.47 | 0.41 | -0.41 | 0 | -0.77* | 0.04 |
| Am3-b | 0.47 | -0.41 | 0.41 | 0 | 0.77* | -0.04 |
| Dia1 | -0.73 | 0.25 | -0.14 | -0.25 | -0.63 | -0.32 |
| dia1 | 0.73 | -0.25 | 0.14 | 0.25 | 0.63 | 0.32 |
| Dia2-a | -0.47 | 0.41 | -0.41 | 0 | -0.77* | 0.04 |
| Dia2-b | 0.47 | -0.41 | 0.41 | 0 | 0.77* | -0.04 |
| Enp-a | 0.74 | -0.29 | 0.18 | 0.29 | 0.56 | 0.31 |
| Enp-b | -0.74 | 0.29 | -0.18 | -0.29 | -0.56 | -0.31 |
| Est-a | 0.89** | -0.14 | 0.11 | 0.58 | 0.19 | 0.67 |
| Est-b | -0.89** | 0.14 | -0.11 | -0.58 | -0.19 | -0.67 |
| Gpd | 0.95** | -0.04 | 0.14 | 0.57 | 0.26 | 0.78* |
| gpd | -0.95** | 0.04 | -0.14 | -0.57 | -0.26 | -0.78* |
| Idh2-a | 0.07 | 0.26 | -0.08 | -0.02 | -0.51 | 0.16 |
| Idh2-b | -0.07 | -0.26 | 0.08 | 0.02 | 0.51 | -0.16 |
| Lap-a | -0.47 | 0.45 | -0.47 | 0.07 | -0.74 | 0.02 |
| Lap-b | 0.47 | -0.45 | 0.47 | -0.07 | 0.74 | -0.02 |
| Mpi-a | -0.61 | 0.09 | -0.18 | 0.27 | 0.05 | -0.58 |
| Mpi-b | 0.61 | -0.09 | 0.18 | -0.27 | -0.05 | 0.58 |
| Pgd1-b | -0.52 | 0.26 | -0.51 | 0.08 | -0.51 | -0.22 |
| Pgd1-c | 0.52 | -0.26 | 0.51 | -0.08 | 0.51 | 0.22 |
| Pgd3-a | -0.53 | 0.20 | -0.20 | 0 | -0.64 | -0.41 |
| Pgd3-b | 0.53 | -0.20 | 0.20 | 0 | 0.64 | 0.41 |
| Pgm1-a | 0.53 | -0.20 | 0.20 | 0 | 0.64 | 0.41 |
| Pgm1-b | -0.53 | 0.20 | -0.20 | 0 | -0.64 | -0.41 |

Appendix XI (1986 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | B4W | L4W | GH | L7W | ONO | BPW |
| Ap-a | 0.20 | 0.62 | 0.10 | 0.62 | 0.41 | 0.41 |
| Ap-c | -0.20 | -0.62 | -0.10 | -0.62 | -0.41 | -0.41 |
| Aco2-a | -0.63 | -0.40 | -0.17 | -0.62 | -0.02 | -0.21 |
| Aco2-b | 0.63 | 0.40 | 0.17 | 0.62 | 0.02 | 0.21 |
| Aco4-a | 0.40 | -0.16 | 0.22 | 0.47 | -0.29 | -0.11 |
| Aco4-b | 0.41 | -0.10 | -0.62 | 0 | 0.62 | 0.62 |
| Aco4-c | -0.67 | 0.20 | 0.38 | -0.36 | -0.31 | -0.45 |
| Am3-s | -0.52 | -0.06 | -0.26 | -0.37 | 0.09 | -0.09 |
| Am3-f | 0.52 | 0.06 | 0.26 | 0.37 | -0.09 | 0.09 |
| Dia1 | -0.04 | 0.11 | -0.72 | -0.25 | 0.05 | -0.02 |
| dia1 | 0.04 | -0.11 | 0.72 | 0.25 | -0.05 | 0.02 |
| Dia2-a | -0.52 | -0.06 | -0.26 | -0.37 | 0.09 | -0.09 |
| Dia2-b | 0.52 | 0.06 | 0.26 | 0.37 | -0.09 | 0.09 |
| Enp-a | 0.11 | -0.07 | 0.71 | 0.27 | 0.05 | 0.34 |
| Enp-b | -0.11 | 0.07 | -0.71 | -0.27 | -0.05 | -0.34 |
| Est-a | -0.32 | 0 | 0.93** | -0.05 | 0.13 | 0.23 |
| Est-b | 0.32 | 0 | -0.93** | 0.05 | -0.13 | -0.23 |
| Gpd | -0.36 | 0.04 | 0.85* | -0.20 | 0.02 | 0.19 |
| gpd | 0.36 | -0.04 | -0.85* | 0.20 | -0.02 | -0.19 |
| Idh2-a | 0.45 | 0.14 | -0.42 | 0.02 | 0.81* | 0.52 |
| Idh2-b | -0.45 | -0.14 | 0.42 | -0.02 | -0.81* | -0.52 |
| Lap-a | -0.56 | -0.16 | -0.22 | -0.41 | 0.07 | -0.12 |
| Lap-b | 0.56 | 0.16 | 0.22 | 0.41 | -0.07 | 0.12 |
| Mpi-a | -0.18 | -0.48 | -0.16 | -0.27 | -0.43 | -0.36 |
| Mpi-b | 0.18 | 0.48 | 0.16 | 0.27 | 0.43 | 0.36 |
| Pgd1-b | -0.51 | -0.35 | 0 | -0.15 | -0.01 | -0.15 |
| Pgd1-c | 0.51 | 0.35 | 0 | 0.15 | 0.01 | 0.15 |
| Pgd3-b | 0.41 | -0.10 | -0.62 | 0 | 0.62 | 0.62 |
| Pgd3-c | -0.41 | 0.10 | 0.62 | 0 | -0.62 | -0.62 |
| Pgm1-a | -0.41 | 0.10 | 0.62 | 0 | -0.62 | -0.62 |
| Pgm1-b | 0.41 | -0.10 | -0.62 | 0 | 0.62 | 0.62 |

Appendix XI (1986 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | FL | CTL | LCL | WCL | LRC | LLL |
| Ap-a | 0.31 | -0.76* | 0.20 | -0.31 | 0.41 | 0.20 |
| Ap-c | -0.31 | 0.76* | -0.20 | 0.31 | -0.41 | -0.20 |
| Aco2-a | 0.32 | 0.55 | -0.44 | 0.30 | -0.93** | -0.37 |
| Aco2-b | -0.32 | -0.55 | 0.44 | -0.30 | 0.93** | 0.37 |
| Aco4-a | -0.59 | 0 | 0.27 | -0.32 | 0.67 | 0.09 |
| Aco4-b | 0.62 | 0.76* | 0.41 | 0.62 | -0.62 | 0.62 |
| Aco4-c | -0.09 | -0.66 | -0.57 | -0.31 | 0.02 | -0.62 |
| Am3-a | 0.47 | 0.28 | -0.26 | 0.28 | -0.82* | -0.22 |
| Am3-b | -0.47 | -0.28 | 0.26 | -0.28 | 0.82* | 0.22 |
| Dia1 | 0.40 | 0.40 | 0.28 | 0.69* | -0.70 | 0.28 |
| dia1 | -0.40 | -0.40 | -0.28 | -0.69 | 0.70 | -0.28 |
| Dia2-a | 0.47 | 0.28 | -0.26 | 0.28 | -0.82* | -0.22 |
| Dia2-b | -0.47 | -0.28 | 0.26 | -0.28 | 0.82* | 0.22 |
| Enp-a | -0.30 | -0.34 | -0.25 | -0.68 | 0.68 | -0.22 |
| Enp-b | 0.30 | 0.34 | 0.25 | 0.68 | -0.68 | 0.22 |
| Est-a | -0.01 | -0.34 | -0.68 | -0.82* | 0.35 | -0.58 |
| Est-b | 0.01 | 0.34 | 0.68 | 0.82 | -0.35 | 0.58 |
| Gpd | -0.07 | -0.40 | -0.68 | -0.71 | 0.34 | -0.57 |
| gpd | 0.07 | 0.40 | 0.68 | 0.71 | -0.34 | 0.57 |
| Idh2-a | 0.69 | 0.30 | 0.32 | 0.46 | -0.34 | 0.63 |
| Idh2-b | -0.69 | -0.30 | -0.32 | -0.46 | 0.34 | -0.63 |
| Lap-a | 0.43 | 0.35 | -0.32 | 0.26 | -0.85* | -0.28 |
| Lap-b | -0.43 | -0.35 | 0.32 | -0.26 | 0.85* | 0.28 |
| Mpi-a | -0.27 | 0.67 | -0.04 | 0.20 | -0.36 | -0.18 |
| Mpi-b | 0.27 | -0.67 | 0.04 | -0.20 | 0.36 | 0.18 |
| Pgd1-b | 0.20 | 0.37 | -0.33 | -0.01 | -0.62 | -0.37 |
| Pgd1-c | -0.20 | -0.37 | 0.33 | 0.01 | 0.62 | 0.37 |
| Pgd3-a | 0.62 | 0.76 | 0.41 | 0.62 | -0.62 | 0.61 |
| Pgd3-b | -0.62 | -0.76 | -0.41 | -0.62 | 0.62 | -0.61 |
| Pgm1-a | -0.62 | -0.76 | -0.41 | -0.62 | 0.62 | -0.61 |
| Pgm1-b | 0.62 | 0.76 | 0.41 | 0.62 | -0.62 | 0.61 |

Appendix XI (1986 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | WLL | LRL | AGW | BGW | NRN | NDW |
| Ap-a | 0.10 | 0.21 | -0.20 | 0.61 | -0.61 | 0.10 |
| Ap-b | -0.10 | -0.21 | 0.20 | -0.61 | 0.61 | -0.10 |
| Aco2-a | 0.75 | -0.92** | 0.63 | -0.70 | 0.89** | 0.63 |
| Aco2-b | -0.75 | 0.92** | -0.63 | 0.70 | -0.89** | -0.63 |
| Aco4-a | -0.81* | 0.69 | -0.76* | 0.36 | -0.49 | -0.76* |
| Aco4-b | 0.62 | -0.42 | 0.61 | 0.20 | 0.61 | 0.62 |
| Aco4-c | -0.08 | -0.16 | -0.04 | -0.45 | -0.16 | 0.04 |
| Am3-a | 0.89** | -0.92** | 0.63 | -0.44 | 0.70 | 0.82* |
| Am3-b | -0.89** | 0.92** | -0.63 | 0.44 | -0.70 | -0.82* |
| Dia1 | 0.83* | -0.70 | 0.78* | -0.14 | 0.68 | 0.95** |
| dia1 | -0.83* | 0.70 | -0.78* | 0.14 | -0.68 | -0.95** |
| Dia2-a | 0.89** | -0.92** | 0.63 | -0.44 | 0.70 | 0.82* |
| Dia2-b | -0.89** | 0.92** | -0.63 | 0.44 | -0.70 | -0.82* |
| Enp-a | -0.79* | 0.69 | -0.77* | 0.20 | -0.67 | -0.94** |
| Enp-b | 0.79* | -0.69 | 0.77* | -0.20 | 0.67 | 0.94** |
| Est-a | -0.43 | 0.24 | -0.63 | -0.16 | -0.45 | -0.76* |
| Est-b | 0.43 | -0.24 | 0.63 | 0.16 | 0.45 | 0.76* |
| Gpd | -0.43 | 0.26 | -0.50 | -0.25 | -0.43 | -0.70 |
| gpd | 0.43 | -0.26 | 0.50 | 0.25 | 0.43 | 0.70 |
| Idh2-a | 0.45 | -0.12 | 0.57 | 0.33 | 0.24 | 0.45 |
| Idh2-b | -0.45 | 0.12 | -0.57 | -0.33 | -0.24 | -0.45 |
| Lap-a | 0.86* | -0.93** | 0.60 | -0.51 | 0.75 | 0.76* |
| Lap-b | -0.86* | 0.93** | -0.60 | 0.51 | -0.75 | -0.76* |
| Mpi-a | 0.04 | -0.32 | 0.04 | -0.40 | 0.53 | 0.04 |
| Mpi-b | -0.04 | 0.32 | -0.04 | 0.40 | -0.53 | -0.04 |
| Pgd1-b | 0.55 | -0.72 | 0.20 | -0.39 | 0.57 | 0.41 |
| Pgd1-c | -0.55 | 0.72 | -0.20 | 0.39 | -0.57 | -0.41 |
| Pgd3-a | 0.62 | -0.42 | 0.61 | 0.20 | 0.61 | 0.62 |
| Pgd3-b | -0.62 | 0.42 | -0.61 | -0.20 | -0.61 | -0.62 |
| Pgm1-a | -0.62 | 0.42 | -0.61 | -0.20 | -0.61 | -0.62 |
| Pgm1-b | 0.62 | -0.42 | 0.61 | 0.20 | 0.61 | 0.62 |

Appendix XI (1986 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | PPP | P1SP | P2SP | P3SP | P4SP | SWP |
| Ap-a | 0.20 | 0.31 | -0.10 | 0 | 0.21 | 0.62 |
| Ap-b | -0.20 | -0.31 | 0.10 | 0 | -0.21 | -0.62 |
| Aco2-a | -0.85* | -0.06 | -0.22 | 0.11 | 0.37 | -0.73 |
| Aco2-b | 0.85* | 0.06 | 0.22 | -0.11 | -0.37 | 0.73 |
| Aco4-a | 0.80* | 0.13 | 0.47 | -0.09 | -0.33 | 0.47 |
| Aco4-b | -0.61 | 0 | 0.21 | -0.20 | 0.64 | -0.41 |
| Aco4-c | -0.07 | 0.09 | -0.55 | 0.26 | -0.30 | 0 |
| Am3-a | -0.85* | 0.24 | -0.17 | 0 | 0.45 | -0.52 |
| Am3-b | 0.85* | -0.24 | 0.17 | 0 | -0.45 | 0.52 |
| Dia1 | -0.82* | 0.59 | 0.23 | -0.43 | 0.24 | -0.58 |
| dia1 | 0.82* | -0.59 | -0.23 | 0.43 | -0.24 | 0.58 |
| Dia2-a | -0.85* | 0.24 | -0.17 | 0 | 0.45 | -0.52 |
| Dia2-b | 0.85* | -0.24 | 0.17 | 0 | -0.45 | 0.52 |
| Enp-a | 0.81* | -0.65 | -0.25 | 0.47 | -0.14 | 0.60 |
| Enp-b | -0.81* | 0.65 | 0.25 | -0.47 | 0.14 | -0.60 |
| Est-a | 0.52 | -0.84* | -0.69 | 0.83* | 0.12 | 0.49 |
| Est-b | -0.52 | 0.84* | 0.69 | -0.83* | -0.12 | -0.49 |
| Gpd | 0.46 | -0.74 | -0.72 | 0.75 | -0.06 | 0.38 |
| gpd | -0.46 | 0.74 | 0.72 | -0.75 | 0.06 | -0.38 |
| Idh2-a | -0.43 | -0.13 | -0.15 | -0.04 | 0.52 | -0.20 |
| Idh2-b | 0.43 | 0.13 | 0.15 | 0.04 | -0.52 | 0.20 |
| Lap-a | -0.84* | 0.16 | -0.17 | 0.04 | 0.45 | -0.56 |
| Lap-b | 0.84* | -0.16 | 0.17 | -0.04 | -0.45 | 0.56 |
| Mpi-a | -0.18 | 0 | 0.36 | -0.13 | -0.05 | -0.40 |
| Mpi-b | 0.18 | 0 | -0.36 | 0.13 | 0.05 | 0.40 |
| Pgd1-b | -0.51 | 0 | -0.01 | 0.10 | 0.44 | -0.33 |
| Pgd1-c | 0.51 | 0 | 0.01 | -0.10 | -0.44 | 0.33 |
| Pgd3-a | -0.61 | 0 | 0.21 | -0.20 | 0.64 | -0.41 |
| Pgd3-b | 0.61 | 0 | -0.21 | 0.20 | -0.64 | 0.41 |
| Pgm1-a | 0.61 | 0 | -0.21 | 0.20 | -0.64 | 0.41 |
| Pgm1-b | -0.61 | 0 | 0.21 | -0.20 | 0.64 | -0.41 |

Appendix XI (1986 Spearman rank coefficients, continued)

| | Quantitative trait | | |
|---|---|---|---|
| Locus | SWH | SPP | ASP |
| Ap-a | 0.31 | 0.61 | 0.41 |
| Ap-b | -0.31 | -0.61 | -0.41 |
| Aco2-a | -0.15 | -0.48 | 0.30 |
| Aco2-b | 0.15 | 0.48 | -0.30 |
| Aco4-a | -0.36 | 0.27 | -0.40 |
| Aco4-b | 0.62 | -0.61 | 0.00 |
| Aco4-c | -0.27 | 0.33 | 0.32 |
| Am3-a | -0.02 | -0.30 | 0.44 |
| Am3-b | 0.02 | 0.30 | -0.44 |
| Dia1 | 0.31 | -0.57 | 0.04 |
| dia1 | -0.31 | 0.57 | -0.04 |
| Dia2-a | -0.02 | -0.30 | 0.44 |
| Dia2-b | 0.02 | 0.30 | -0.44 |
| Enp-a | -0.25 | 0.56 | 0 |
| Enp-b | 0.25 | -0.56 | 0 |
| Est-a | -0.43 | 0.63 | 0.43 |
| Est-b | 0.43 | -0.63 | -0.43 |
| Gpd | -0.36 | 0.54 | 0.32 |
| gpd | 0.36 | -0.54 | -0.32 |
| Idh2-a | 0.90** | -0.37 | 0.06 |
| Idh2-b | -0.90** | 0.37 | -0.06 |
| Lap-a | -0.09 | -0.32 | 0.43 |
| Lap-b | 0.09 | 0.32 | -0.43 |
| Mpi-a | -0.48 | -0.40 | -0.22 |
| Mpi-b | 0.48 | 0.40 | 0.22 |
| Pgd1-b | -0.35 | -0.12 | 0.39 |
| Pgd1-c | 0.35 | 0.12 | -0.39 |
| Pgd3-a | 0.62 | -0.61 | 0 |
| Pgd3-b | -0.62 | 0.61 | 0 |
| Pgm1-a | -0.62 | 0.61 | 0 |
| Pgm1-b | 0.62 | -0.61 | 0 |

Appendix XI (continued)

B. 1987 Spearman rank correlation coefficients between isozyme loci and quantiative traits

| | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| Locus | DTG | DTF | DFP | DDP | LDP | LSN |
| Ap-a | 0.32 | -0.20 | 0 | -0.52 | 0.41 | 0 |
| Ap-c | -0.32 | 0.20 | 0 | 0.52 | -0.41 | 0 |
| Aco2-a | -0.50 | 0.37 | -0.30 | 0.06 | -0.93** | 0.11 |
| Aco2-b | 0.50 | -0.37 | 0.30 | -0.06 | 0.93 | -0.11 |
| Aco4-a | 0.02 | -0.13 | 0.27 | 0.20 | 0.67 | -0.13 |
| Aco4-b | -0.21 | 0 | 0.20 | -0.10 | -0.51 | 0.20 |
| Aco4-c | 0.18 | 0.12 | -0.39 | -0.08 | -0.07 | -0.06 |
| Am3-s | -0.56 | 0.15 | -0.18 | -0.02 | -0.82* | -0.07 |
| Am3-b | 0.56 | -0.15 | 0.18 | 0.02 | 0.82* | 0.07 |
| Dia1 | -0.65 | -0.36 | 0.32 | 0.33 | -0.65 | -0.43 |
| dia1 | 0.65 | 0.36 | -0.32 | -0.33 | 0.65 | 0.43 |
| Dia2-a | -0.56 | 0.15 | -0.18 | -0.02 | -0.82* | -0.07 |
| Dia2-b | 0.56 | -0.15 | 0.18 | 0.02 | 0.82* | 0.07 |
| Enp-a | 0.66 | 0.40 | -0.34 | -0.38 | 0.63 | 0.49 |
| Enp-b | -0.66 | -0.40 | 0.34 | 0.38 | -0.63 | -0.49 |
| Est-a | 0.51 | 0.76 | -0.77* | -0.52 | 0.23 | 0.63 |
| Est-b | -0.51 | -0.76 | 0.77* | 0.52 | -0.23 | -0.63 |
| Gpd | 0.56 | 0.64 | -0.71 | -0.40 | 0.22 | 0.54 |
| gpd | -0.56 | -0.64 | 0.71 | 0.40 | -0.22 | -0.54 |
| Idh2-a | 0.40 | 0.04 | 0.04 | -0.41 | -0.25 | 0.45 |
| Idh2-b | -0.40 | -0.04 | -0.04 | 0.41 | 0.25 | -0.45 |
| Lap-a | -0.57 | 0.22 | -0.22 | -0.02 | -0.85* | -0.02 |
| Lap-b | 0.57 | -0.22 | 0.22 | 0.02 | 0.85* | 0.02 |
| Mpi-a | -0.70 | 0.04 | 0.13 | 0.52 | -0.36 | -0.27 |
| Mpi-b | 0.70 | -0.04 | -0.13 | -0.52 | 0.36 | 0.27 |
| Pgd1-b | -0.57 | 0.33 | -0.23 | -0.08 | -0.61 | 0.06 |
| Pgd1-c | 0.57 | -0.33 | 0.23 | 0.08 | 0.61 | -0.06 |
| Pgd3-a | -0.21 | 0 | 0.20 | -0.10 | -0.51 | 0.20 |
| Pgd3-b | 0.21 | 0 | 0.20 | 0.10 | 0.51 | -0.20 |
| Pgm1-a | 0.21 | 0 | -0.20 | 0.10 | 0.51 | -0.20 |
| Pgm1-b | -0.21 | 0 | 0.20 | -0.10 | -0.51 | 0.20 |

Appendix XI (1987 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | B4W | L4W | GH | L7W | ONO | BPW |
| Ap-a | -0.10 | -0.11 | 0 | -0.20 | 0.62 | -0.21 |
| Ap-c | 0.10 | 0.11 | 0 | 0.20 | -0.62 | 0.21 |
| Aco2-a | 0.32 | 0.08 | -0.21 | -0.33 | 0.06 | -0.32 |
| Aco2-b | -0.32 | -0.08 | 0.21 | 0.33 | -0.06 | 0.32 |
| Aco4-a | 0.04 | 0.20 | 0.22 | 0.36 | -0.31 | 0.16 |
| Aco4-b | 0.41 | -0.11 | -0.62 | -0.41 | 0.41 | 0.41 |
| Aco4-c | -0.39 | -0.04 | 0.38 | 0.08 | -0.11 | -0.48 |
| Am3-s | 0.32 | -0.02 | -0.37 | -0.44 | 0.28 | -0.52 |
| Am3-f | -0.32 | 0.02 | 0.37 | 0.44 | -0.28 | 0.52 |
| Dia1 | 0.16 | -0.35 | -0.81* | -0.46 | -0.09 | -0.36 |
| dia1 | -0.16 | 0.35 | 0.81* | 0.46 | 0.09 | 0.36 |
| Dia2-a | 0.32 | -0.02 | -0.37 | -0.44 | 0.28 | -0.52 |
| Dia2-b | -0.32 | 0.02 | 0.37 | 0.44 | -0.28 | 0.52 |
| Enp-a | -0.07 | 0.41 | 0.80* | 0.49 | -0.01 | 0.45 |
| Enp-b | 0.07 | -0.41 | -0.80* | -0.49 | 0.01 | -0.45 |
| Est-a | 0.18 | 0.68 | 0.98** | 0.56 | 0.17 | 0.27 |
| Est-b | -0.18 | -0.68 | -0.98** | -0.56 | -0.17 | -0.27 |
| Gpd | -0.02 | 0.53 | 0.94** | 0.57 | 0.02 | 0.27 |
| gpd | 0.02 | -0.53 | -0.94** | -0.57 | -0.02 | -0.27 |
| Idh2-a | 0.03 | -0.23 | -0.32 | -0.31 | 0.52 | 0.67 |
| Idh2-b | -0.03 | 0.23 | 0.32 | 0.31 | -0.52 | -0.67 |
| Lap-a | 0.35 | 0.03 | -0.32 | -0.43 | 0.24 | -0.50 |
| Lap-b | -0.35 | -0.03 | 0.32 | 0.43 | -0.24 | 0.50 |
| Mpi-a | 0.36 | 0.17 | -0.20 | 0.04 | -0.40 | -0.22 |
| Mpi-b | -0.36 | -0.17 | 0.20 | -0.04 | 0.40 | 0.22 |
| Pgd1-b | 0.48 | 0.23 | -0.13 | -0.35 | 0.23 | -0.51 |
| Pgd1-c | -0.48 | -0.23 | 0.13 | 0.35 | -0.23 | 0.51 |
| Pgd3-a | 0.41 | -0.11 | -0.62 | -0.41 | 0.41 | 0.41 |
| Pgd3-b | -0.41 | 0.11 | 0.62 | 0.41 | -0.41 | -0.41 |
| Pgm1-a | -0.41 | 0.11 | 0.62 | 0.41 | -0.41 | -0.41 |
| Pgm1-b | 0.41 | -0.11 | -0.62 | -0.41 | 0.41 | 0.41 |

Appendix XI (1987 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | FL | CTL | LCL | WCL | LRC | LLL |
| Ap-a | 0 | 0.32 | 0.21 | -0.10 | 0.31 | 0.62 |
| Ap-c | 0 | -0.32 | -0.21 | 0.10 | -0.31 | -0.62 |
| Aco2-a | -0.73 | -0.16 | -0.62 | 0.84* | -0.90** | -0.93** |
| Aco2-b | 0.73 | 0.16 | 0.62 | -0.84* | 0.90** | 0.93** |
| Aco4-a | 0.47 | 0.23 | 0.47 | -0.81* | 0.74 | 0.36 |
| Aco4-b | -0.21 | 0.32 | -0.62 | 0.62 | -0.62 | -0.31 |
| Aco4-c | -0.18 | -0.44 | 0.17 | 0.08 | -0.03 | -0.01 |
| Am3-a | -0.90** | -0.12 | -0.50 | 0.90** | -0.84* | -0.75 |
| Am3-b | 0.90** | 0.12 | 0.50 | -0.90** | 0.84* | 0.75 |
| Dia1 | -0.83* | -0.32 | -0.27 | 0.86* | -0.77* | -0.52 |
| dia1 | 0.83* | 0.32 | 0.27 | -0.86* | 0.77* | 0.52 |
| Dia2-a | -0.90** | -0.12 | -0.50 | 0.90** | -0.84* | -0.75 |
| Dia2-b | 0.90* | 0.12 | 0.50 | -0.90** | 0.84* | 0.75 |
| Enp-a | 0.87* | 0.41 | 0.18 | -0.85* | 0.76* | 0.53 |
| Enp-b | -0.87* | -0.41 | -0.18 | 0.85* | -0.76* | -0.53 |
| Est-a | 0.62 | 0.41 | -0.18 | -0.56 | 0.47 | 0.16 |
| Est-b | -0.62 | -0.41 | 0.18 | 0.65 | -0.47 | -0.16 |
| Gpd | 0.63 | 0.17 | -0.09 | -0.52 | 0.43 | 0.20 |
| gpd | -0.63 | -0.17 | 0.09 | 0.52 | -0.43 | -0.20 |
| Idh2-a | 0.24 | 0.30 | -0.51 | 0.42 | -0.39 | 0.17 |
| Idh2-b | -0.24 | -0.30 | 0.51 | -0.42 | 0.39 | -0.17 |
| Lap-a | -0.88** | -0.09 | -0.53 | 0.88** | -0.85* | -0.81* |
| Lap-b | 0.88** | 0.09 | 0.53 | -0.88** | 0.85* | 0.81* |
| Mpi-a | -0.40 | -0.14 | -0.11 | 0.16 | -0.27 | -0.67 |
| Mpi-b | 0.40 | 0.14 | 0.11 | -0.16 | 0.27 | 0.67 |
| Pgd1-b | -0.74 | 0.17 | -0.43 | 0.56 | -0.57 | -0.76 |
| Pgd1-c | 0.74 | -0.17 | 0.43 | -0.56 | 0.57 | 0.76 |
| Pgd3-a | -0.21 | 0.32 | -0.62 | 0.62 | -0.62 | -0.31 |
| Pgd3-b | 0.21 | -0.32 | 0.62 | -0.62 | 0.62 | 0.31 |
| Pgm1-a | 0.21 | -0.32 | 0.62 | -0.62 | 0.62 | 0.31 |
| Pgm1-b | -0.21 | 0.32 | -0.62 | 0.62 | -0.62 | -0.31 |

Appendix XI (1987 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | WLL | LRL | AGW | BGW | NRN | NDW |
| Ap-a | -0.41 | 0.41 | -0.20 | -0.20 | -0.41 | -0.20 |
| Ap-b | 0.41 | -0.41 | 0.20 | 0.20 | 0.41 | 0.20 |
| Aco2-a | 0.65 | -0.93** | -0.33 | -0.26 | 0.07 | -0.07 |
| Aco2-b | -0.65 | 0.93** | 0.33 | 0.26 | -0.07 | 0.07 |
| Aco4-a | -0.27 | 0.47 | 0.36 | 0.58 | 0.53 | 0.49 |
| Aco4-b | 0.62 | -0.62 | 0.61 | 0.41 | -0.20 | 0.61 |
| Aco4-c | -0.32 | 0.17 | -0.80* | -0.79 | -0.24 | -0.91** |
| Am3-a | 0.41 | -0.80* | -0.44 | -0.44 | 0.04 | -0.11 |
| Am3-b | -0.41 | 0.80* | 0.44 | 0.44 | -0.04 | 0.11 |
| Dia1 | 0.18 | -0.56 | -0.11 | -0.43 | 0.18 | 0.07 |
| dia1 | -0.18 | 0.56 | 0.11 | 0.43 | -0.18 | -0.07 |
| Dia2-a | 0.41 | -0.80* | -0.44 | -0.44 | 0.04 | -0.11 |
| Dia2-b | -0.41 | 0.80* | 0.44 | 0.44 | -0.04 | 0.11 |
| Enp-a | -0.14 | 0.52 | 0.22 | 0.52 | -0.23 | 0.04 |
| Enp-b | 0.14 | -0.52 | -0.22 | -0.52 | 0.23 | -0.04 |
| Est-a | 0.01 | -0.15 | -0.07 | 0.31 | -0.38 | -0.11 |
| Est-b | -0.01 | 0.15 | 0.07 | -0.31 | 0.38 | 0.11 |
| Gpd | -0.09 | 0.25 | -0.14 | 0.11 | -0.43 | -0.32 |
| gpd | 0.09 | -0.25 | 0.14 | -0.11 | 0.43 | 0.32 |
| Idh2-a | 0.41 | -0.20 | 0.57 | 0.26 | -0.79* | 0.16 |
| Idh2-b | -0.41 | 0.20 | -0.57 | -0.26 | 0.79* | -0.16 |
| Lap-a | 0.49 | -0.86* | -0.43 | -0.37 | 0.07 | -0.07 |
| Lap-b | -0.49 | 0.86* | 0.43 | 0.37 | -0.07 | 0.07 |
| Mpi-a | 0.29 | -0.49 | 0.04 | 0.18 | 0.76* | 0.44 |
| Mpi-b | -0.29 | 0.49 | -0.04 | -0.18 | -0.76* | -0.44 |
| Pgd1-b | 0.53 | -0.80* | -0.35 | -0.04 | 0.32 | 0.18 |
| Pgd1-c | -0.53 | 0.80* | 0.35 | 0.04 | -0.32 | -0.18 |
| Pgd3-a | 0.62 | -0.62 | 0.61 | 0.41 | -0.20 | 0.61 |
| Pgd3-b | -0.62 | 0.62 | -0.61 | -0.41 | 0.20 | -0.61 |
| Pgm1-a | -0.62 | 0.62 | -0.61 | -0.41 | 0.20 | -0.61 |
| Pgm1-b | 0.62 | 0.62 | 0.61 | 0.41 | -0.20 | 0.61 |

Appendix XI (1987 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | | | | |
|---|---|---|---|---|---|---|
| | PPP | P1SP | P2SP | P3SP | P4SP | SWP |
| Ap-a | 0.41 | 0.20 | 0 | 0 | -0.20 | 0.21 |
| Ap-b | -0.41 | -0.20 | 0 | 0 | 0.20 | -0.21 |
| Aco2-a | -0.96** | 0.33 | -0.30 | -0.07 | 0.26 | -0.71 |
| Aco2-b | 0.96** | -0.33 | 0.30 | 0.07 | -0.26 | 0.71 |
| Aco4-a | 0.53 | -0.27 | 0.27 | 0.04 | -0.13 | 0.13 |
| Aco4-b | -0.61 | 0.41 | 0.20 | -0.20 | 0.20 | 0 |
| Aco4-c | 0.12 | -0.16 | -0.39 | 0.16 | -0.06 | -0.12 |
| Am3-a | -0.85* | 0.59 | -0.18 | -0.26 | 0 | -0.60 |
| Am3-b | 0.85* | -0.59 | 0.18 | 0.26 | 0 | 0.60 |
| Dia1 | -0.64 | 0.82* | 0.32 | -0.68 | -0.39 | -0.09 |
| dia1 | 0.64 | -0.82* | -0.32 | 0.68 | 0.39 | 0.09 |
| Dia2-a | -0.85* | 0.59 | -0.18 | -0.26 | 0 | -0.60 |
| Dia2-b | 0.85* | -0.59 | 0.18 | 0.26 | 0 | 0.60 |
| Enp-a | 0.61 | -0.85* | -0.34 | 0.72 | 0.45 | 0.12 |
| Enp-b | -0.61 | 0.85* | 0.34 | -0.72 | -0.45 | -0.12 |
| Est-a | 0.25 | -0.92** | -0.77* | 0.94** | 0.67 | -0.21 |
| Est-b | -0.25 | 0.92** | 0.77* | -0.94** | -0.67 | 0.21 |
| Gpd | 0.32 | -0.93** | -0.71 | 0.86* | 0.57 | -0.05 |
| gpd | -0.32 | 0.93** | 0.71 | -0.86* | -0.57 | 0.05 |
| Idh2-a | -0.20 | 0.04 | 0.04 | 0.10 | 0.35 | 0.38 |
| Idh2-b | 0.20 | -0.04 | -0.04 | -0.10 | -0.35 | -0.38 |
| Lap-a | -0.90** | 0.54 | -0.22 | -0.21 | 0.07 | -0.68 |
| Lap-b | 0.90** | -0.54 | 0.22 | 0.21 | -0.07 | 0.68 |
| Mpi-a | -0.50 | 0.22 | 0.13 | -0.27 | -0.09 | -0.45 |
| Mpi-b | 0.50 | -0.22 | -0.13 | 0.27 | 0.09 | 0.45 |
| Pgd1-b | -0.79* | 0.43 | -0.24 | -0.08 | 0.16 | -0.87** |
| Pgd1-c | 0.79* | -0.43 | 0.24 | 0.08 | -0.16 | 0.87** |
| Pgd3-a | -0.61 | 0.41 | 0.20 | -0.20 | 0.20 | 0 |
| Pgd3-b | 0.61 | -0.41 | -0.20 | 0.20 | -0.20 | 0 |
| Pgm1-a | 0.61 | -0.41 | -0.20 | 0.20 | -0.20 | 0 |
| Pgm1-b | -0.61 | 0.41 | 0.20 | -0.20 | 0.20 | 0 |

Appendix XI (1987 Spearman rank coefficients, continued)

| Locus | Quantitative trait | | |
|---|---|---|---|
| | SWH | SPP | ASP |
| Ap-a | 0.20 | 0 | 0 |
| Ap-b | -0.20 | 0 | 0 |
| Aco2-a | 0.11 | -0.70 | 0.06 |
| Aco2-b | -0.11 | 0.70 | -0.06 |
| Aco4-a | -0.27 | 0.40 | -0.31 |
| Aco4-b | 0.61 | -0.61 | -0.21 |
| Aco4-c | -0.33 | 0.22 | 0.43 |
| Am3-a | 0.30 | -0.74 | -0.09 |
| Am3-b | -0.30 | 0.74 | 0.09 |
| Dia1 | 0.64 | -0.57 | -0.45 |
| dia1 | -0.64 | 0.57 | 0.45 |
| Dia2-a | 0.30 | -0.74 | 0.09 |
| Dia2-b | -0.30 | 0.74 | -0.09 |
| Enp-a | -0.63 | 0.56 | 0.48 |
| Enp-b | 0.63 | -0.56 | -0.48 |
| Est-a | -0.85* | 0.41 | 0.81* |
| Est-b | 0.85* | -0.41 | -0.81* |
| Gpd | -0.82* | 0.54 | 0.85* |
| gpd | 0.82* | -0.54 | -0.85* |
| Idh2-a | 0.49 | -0.26 | 0.24 |
| Idh2-b | -0.49 | 0.26 | -0.24 |
| Lap-a | 0.24 | -0.77* | -0.08 |
| Lap-b | -0.24 | 0.77* | 0.08 |
| Mpi-a | -0.04 | -0.27 | -0.40 |
| Mpi-b | 0.04 | 0.27 | 0.40 |
| Pgd1-b | 0.08 | -0.77* | -0.19 |
| Pgd1-c | -0.08 | 0.77* | 0.19 |
| Pgd3-a | 0.61 | -0.61 | -0.21 |
| Pgd3-b | -0.61 | 0.61 | 0.21 |
| Pgm1-a | -0.61 | 0.61 | 0.21 |
| Pgm1-b | 0.61 | -0.61 | -0.21 |